

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 8, 2005, 16:53:04 ; Search time 2363 Seconds
(without alignments)
10129.867 Million cell updates/sec

Title: US-09-925-824A-1
Sequence: 1 gctaatgctgtgtacacg.....gaataccttattgtgcag 494

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues
Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:
1: gb_ba:*
2: gb_hcg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pac:*
7: gb_pl:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	494	100.0	494	6	AX472612	AX472612 Sequence
2	494	100.0	23775	1	BACYP1A	L47709 Bacillus su
3	494	100.0	201375	1	BSUB0012	Z99115 Bacillus su
4	485	98.2	2363	6	AX100815	AX100815 Sequence
5	258	52.2	7281	6	AX100850	AX100850 Sequence
6	257	52.0	6725	6	AX572874	AX572874 Sequence
7	257	52.0	6725	6	AX573226	AX573226 Sequence
8	257	52.0	8503	6	AX100849	AX100849 Sequence
9	257	52.0	10801	6	AX100832	AX100832 Sequence
10	244	49.4	831	6	AX100779	AX100779 Sequence
11	229	46.4	110000	1	AE017333_23	Continuation (24 o
12	229	46.4	110000	1	CP000002_23	Continuation (24 o
13	216.2	43.8	6886	6	AX572872	AX572872 Sequence
14	216.2	43.8	6886	6	AX573224	AX573224 Sequence
15	203	41.1	6540	1	BACJOJC	L38424 Bacillus su
16	161.8	32.8	294300	1	AE017269	AE017269 Bacillus
17	160.2	32.4	110000	1	AE017355_14	Continuation (15 o
18	160.2	32.4	304680	1	AE017002	AE017002 Bacillus
19	158.6	32.1	110000	1	AE017225_14	Continuation (15 o

20	158.6	32.1	110000	1	AE017334_14	Continuation (15 o
21	158.6	32.1	110000	1	CP000001_14	Continuation (15 o
22	158.6	32.1	110000	1	CP000001_15	Continuation (16 o
23	158.6	32.1	290525	1	AE017029	AE017029 Bacillus
24	151.2	30.6	834	6	AX432591	AX432591 Sequence
25	137	27.7	1613	6	AX416270	AX416270 Sequence
26	137	27.7	313450	1	AL596170	AL596170 Listeria
27	137	27.7	313450	6	AX413016	AX413016 Sequence
28	137	27.7	349980	6	AX417046	AX417046 Sequence
29	133.8	27.1	291954	1	AE017328	AE017328 Listeria
30	133.8	27.1	347050	1	AL591981	AL591981 Listeria
31	133.8	27.1	349980	6	AX641670	AX641670 Sequence
32	117.2	23.7	1880	1	AF270386	AF270386 Staphyloc
33	117.2	23.7	1880	6	AR486340	AR486340 Sequence
34	117.2	23.7	1880	6	AX145704	AX145704 Sequence
35	117.2	23.7	300275	1	AE016751	AE016751 Staphyloc
36	114.4	23.2	300550	1	AP001512	AP001512 Bacillus
37	103.2	20.9	792	6	AR484290	AR484290 Sequence
38	103.2	20.9	792	6	AX142695	AX142695 Sequence
39	99.6	20.2	4549	6	AR535612	AR535612 Sequence
40	99.6	20.2	4549	6	BS571856_27	Continuation (28 o
41	99.6	20.2	110000	1	BS571857_26	Continuation (27 o
42	99.6	20.2	110000	1	AP003137	AP003137 Staphyloc
43	99.6	20.2	297850	1	AP004830	AP004830 Staphyloc
44	99.6	20.2	301250	1	AP003365	AP003365 Staphyloc
45	98	19.8	342350	1	AP003365	AP003365 Staphyloc

ALIGNMENTS

RESULT 1
LOCUS AX472612 494 bp DNA linear PAT 09-AUG-2002
DEFINITION Sequence 1 from Patent WO0214559.
ACCESSION AX472612
VERSION AX472612.1 GI:22207504
KEYWORDS
SOURCE
ORGANISM Bacillus subtilis
Bacillus subtilis
Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
REFERENCE
1 Murphy, C.
High throughput screen for inhibitors of the folate biosynthetic
pathway in bacteria
Patent: WO 0214559-A 1 21-FEB-2002;
JOURNAL Millennium Pharmaceuticals, Inc. (US)
Milleium Pharmaceuticals, Inc. (US)
FEATURES
source
1..494
/organism="Bacillus subtilis"
/mol_type="unassigned DNA"
/db_xref="taxon:1423"

ORIGIN

Query Match 100.0%; Score 494; DB 6; Length 494;
Best Local Similarity 100.0%; Pred. No. 4.3e-123; Indels 0; Gaps 0;
Matches 494; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1 GCTAATGTTGGTACCAAGCCGTTGATTTGGTATTAATCTTCATTGGCAGTATGCGCTG 60
2 GCTAATGTTGGTACCAAGCCGTTGATTTGGTATTAATCTTCATTGGCAGTATGCGCTG 60
3 GCTAATGTTGGTACCAAGCCGTTGATTTGGTATTAATCTTCATTGGCAGTATGCGCTG 60
4 GCTAATGTTGGTACCAAGCCGTTGATTTGGTATTAATCTTCATTGGCAGTATGCGCTG 60
5 GCTAATGTTGGTACCAAGCCGTTGATTTGGTATTAATCTTCATTGGCAGTATGCGCTG 60
6 GCTAATGTTGGTACCAAGCCGTTGATTTGGTATTAATCTTCATTGGCAGTATGCGCTG 60
7 GCTAATGTTGGTACCAAGCCGTTGATTTGGTATTAATCTTCATTGGCAGTATGCGCTG 60
8 GCTAATGTTGGTACCAAGCCGTTGATTTGGTATTAATCTTCATTGGCAGTATGCGCTG 60
9 GCTAATGTTGGTACCAAGCCGTTGATTTGGTATTAATCTTCATTGGCAGTATGCGCTG 60
10 GCTAATGTTGGTACCAAGCCGTTGATTTGGTATTAATCTTCATTGGCAGTATGCGCTG 60
11 GCTAATGTTGGTACCAAGCCGTTGATTTGGTATTAATCTTCATTGGCAGTATGCGCTG 60
12 GCTAATGTTGGTACCAAGCCGTTGATTTGGTATTAATCTTCATTGGCAGTATGCGCTG 60
13 GCTAATGTTGGTACCAAGCCGTTGATTTGGTATTAATCTTCATTGGCAGTATGCGCTG 60
14 GCTAATGTTGGTACCAAGCCGTTGATTTGGTATTAATCTTCATTGGCAGTATGCGCTG 60
15 GCTAATGTTGGTACCAAGCCGTTGATTTGGTATTAATCTTCATTGGCAGTATGCGCTG 60
16 GCTAATGTTGGTACCAAGCCGTTGATTTGGTATTAATCTTCATTGGCAGTATGCGCTG 60
17 GCTAATGTTGGTACCAAGCCGTTGATTTGGTATTAATCTTCATTGGCAGTATGCGCTG 60
18 GCTAATGTTGGTACCAAGCCGTTGATTTGGTATTAATCTTCATTGGCAGTATGCGCTG 60
19 GCTAATGTTGGTACCAAGCCGTTGATTTGGTATTAATCTTCATTGGCAGTATGCGCTG 60
20 GCTAATGTTGGTACCAAGCCGTTGATTTGGTATTAATCTTCATTGGCAGTATGCGCTG 60
21 GCTAATGTTGGTACCAAGCCGTTGATTTGGTATTAATCTTCATTGGCAGTATGCGCTG 60
22 GCTAATGTTGGTACCAAGCCGTTGATTTGGTATTAATCTTCATTGGCAGTATGCGCTG 60
23 GCTAATGTTGGTACCAAGCCGTTGATTTGGTATTAATCTTCATTGGCAGTATGCGCTG 60
24 GCTAATGTTGGTACCAAGCCGTTGATTTGGTATTAATCTTCATTGGCAGTATGCGCTG 60
25 GCTAATGTTGGTACCAAGCCGTTGATTTGGTATTAATCTTCATTGGCAGTATGCGCTG 60
26 GCTAATGTTGGTACCAAGCCGTTGATTTGGTATTAATCTTCATTGGCAGTATGCGCTG 60
27 GCTAATGTTGGTACCAAGCCGTTGATTTGGTATTAATCTTCATTGGCAGTATGCGCTG 60
28 GCTAATGTTGGTACCAAGCCGTTGATTTGGTATTAATCTTCATTGGCAGTATGCGCTG 60
29 GCTAATGTTGGTACCAAGCCGTTGATTTGGTATTAATCTTCATTGGCAGTATGCGCTG 60
30 GCTAATGTTGGTACCAAGCCGTTGATTTGGTATTAATCTTCATTGGCAGTATGCGCTG 60
31 GCTAATGTTGGTACCAAGCCGTTGATTTGGTATTAATCTTCATTGGCAGTATGCGCTG 60
32 GCTAATGTTGGTACCAAGCCGTTGATTTGGTATTAATCTTCATTGGCAGTATGCGCTG 60
33 GCTAATGTTGGTACCAAGCCGTTGATTTGGTATTAATCTTCATTGGCAGTATGCGCTG 60
34 GCTAATGTTGGTACCAAGCCGTTGATTTGGTATTAATCTTCATTGGCAGTATGCGCTG 60
35 GCTAATGTTGGTACCAAGCCGTTGATTTGGTATTAATCTTCATTGGCAGTATGCGCTG 60
36 GCTAATGTTGGTACCAAGCCGTTGATTTGGTATTAATCTTCATTGGCAGTATGCGCTG 60
37 GCTAATGTTGGTACCAAGCCGTTGATTTGGTATTAATCTTCATTGGCAGTATGCGCTG 60
38 GCTAATGTTGGTACCAAGCCGTTGATTTGGTATTAATCTTCATTGGCAGTATGCGCTG 60
39 GCTAATGTTGGTACCAAGCCGTTGATTTGGTATTAATCTTCATTGGCAGTATGCGCTG 60
40 GCTAATGTTGGTACCAAGCCGTTGATTTGGTATTAATCTTCATTGGCAGTATGCGCTG 60
41 GCTAATGTTGGTACCAAGCCGTTGATTTGGTATTAATCTTCATTGGCAGTATGCGCTG 60
42 GCTAATGTTGGTACCAAGCCGTTGATTTGGTATTAATCTTCATTGGCAGTATGCGCTG 60
43 GCTAATGTTGGTACCAAGCCGTTGATTTGGTATTAATCTTCATTGGCAGTATGCGCTG 60
44 GCTAATGTTGGTACCAAGCCGTTGATTTGGTATTAATCTTCATTGGCAGTATGCGCTG 60
45 GCTAATGTTGGTACCAAGCCGTTGATTTGGTATTAATCTTCATTGGCAGTATGCGCTG 60

Qy 241 AGAGAAAACATGAAAAAAGAACTGATTTTCTMAAAATGAAGAGCTGGAAGAACCGAT 300
| | | | |
Db 241 AGGAGAAAACATGAAAAAAGAACTGATTTTCTMAAAATGAAGAGCTGGAAGAACCGAT 300
Qy 301 TGTCACTGTCGACCGCTTATGATTTATCCGCGACGCTTAACTTCTGAAACAGCGGAGTTGA 360
| | | | |
Db 301 TGTCACTGTCGACCGCTTATGATTTATCCGCGACGCTTAACTTCTGAAACAGCGGAGTTGA 360
Qy 361 CATGATTTTACCTGCTGATTTCACTTGGAATGTCGTCTCTCGGCTTGAATTTCACTGTCGG 420
| | | | |
Db 361 CATGATTTTACCTGCTGATTTCACTTGGAATGTCGTCTCTCGGCTTGAATTTCACTGTCGG 420
Qy 421 TGTGACAGTTGCGGACATGATTCATCATACAAAGACCGTTAAAGGGGTGGCCGAATAC 480
| | | | |
Db 421 TGTGACAGTTGCGGACATGATTCATCATACAAAGACCGTTAAAGGGGTGGCCGAATAC 480
Qy 481 CTTTATTGTGACAG 494
| | | | |
Db 481 CTTTATTGTGACAG 494

RESULT 2

BACYP1A 23775 bp DNA linear BCT 31-JAN-2003
LOCUS
DEFINITION Bacillus subtilis clone YAC15-6B ypiABF genes, qcrABC genes,
ypjABCFDEFGHI genes, bira gene, panBCD genes, dng gene, ypmB gene,
aapB gene, aens gene, dnaD gene, nch gene and ypoC gene, complete
cds.

ACCESSION L47709 GI:1146223
VERSION L47709.1
KEYWORDS
SOURCE

*ORGANISM

Bacillus subtilis
Bacillus subtilis
Bacteria; Firmicutes; Bacilliales; Bacillaceae; Bacillus.
1 (bases 1 to 23775)
Hemmer,D., Gollnick,P. and Moir,A.
AUTHORS
TITLE Analysis of an 18 kilobase pair region of the Bacillus subtilis
chromosome containing the mtr and gerc operons and the aro-try-aro
supraperon

Proc. Int. Symp. Genet. Ind. Microorg. 6, 657-665 (1990)
2 (bases 1 to 23775)
Sorokin,A., Azevedo,V., Zumbstein,E., Galleron,N., Ehrlich,S.D. and
Serror,P.

JOURNAL Sequence analysis of the Bacillus subtilis chromosome region
REFERENCE between the sera and kdg loci cloned in a yeast artificial
AUTHORS chromosome
TITLE Microbiology 142 (Pt 8), 2005-2016 (1996)
JOURNAL MEDLINE 96349105
PUBMED 8760912

FEATURES
source
Location/Qualifiers
1..23775
/organism="Bacillus subtilis"
/mol_type="genomic DNA"
/strain="Marburg"
/sub_strain="168"
/db_xref="taxon:1423"
/map="202 degrees"
/clone="YAC15-6B"
/clone_lib="chromosome YAC library of Pascale Serror"
/note="bases 57126 to 80902 of the YAC15-6B clone"
1..33
/gene="aroE"
1..33
/gene="aroE"
/standard_name="transcription terminator"
/note="putative"
/citation=[2]
348..1619
/gene="ypjA"
348..1619
/gene="ypjA"
/function="hypothetical"

gene
CDS

/note="putative"
/citation=[2]
/codon_start=1
/transl_table=11
/protein_id="AAB38432.1"
/db_xref="GI:1146224"
/translation="MNVLIQEAIKLVEAGETEGKQLNTLSKAERQLHDEDEKALIAAQLY
EMDIVERKALISLIDLDLYPNETLNFYAEILIDIDEERKALAVLETTIPETPSYE
SLIMADLYOMOGLEFVSEBOKLFOAKSILIDNEPVIDPALDELFAOGAVAKAVQYFKT
TAEOSRISGVVNHORLAESLSASGEREDAIPWYERKAVDENPDNTIFGYGFTALOG
LVKTAIKQLSDLEKLDPSYTSUYPMLASKSFEAGMTBEAKTKAGEGRYBYNKEPL
YAKMAKIKSESEKKLLQELALDPGFVEALHTLLAVTHKEEDYDQIIDLQEVNS
YAEEDPKYNNYLAAYTELEQYERAKQSFPAAYLAHYREDRDLFEYASFLLEGLOKE
ALPLIKVLMDPANBELBETIIRIEDEFER"
1691..2230
/gene="ypjB"
1691..2230
/gene="ypjB"
/function="hypothetical"
/note="putative"
/citation=[2]
/codon_start=1
/transl_table=11
/protein_id="AAB38433.1"
/db_xref="GI:1146225"
/translation="MQTPYVNEKQDFIRFWFLNHQYLKRECYWIIINYLMSHDLMEK
VHFVEOAEFCPEPRGIMSTHCVEVPFPEYKENTMTDAEKSFDIRLNKOQDFICLN
FRSAYSGPEYAAVLESNPHI PKNLFEKKOQGLAEOILHHAISTFQREKLIKIDIDL
DRHDKERFEOLSQNLQLT"
2285..2731
/gene="ypjF"
2285..2731
/gene="ypjF"
/function="hypothetical"
/note="putative"
/citation=[2]
/codon_start=1
/transl_table=11
/protein_id="AAB38434.1"
/db_xref="GI:1146226"
/translation="MRWRITDADQYIDQAKYIDTAVYPLINIRVNNHFRMAARKGERT
QLUSEERQLKGVLYLPPYTVDRNEITVQGLKQREILTFEPHVLITSDSWR
AEDALGMIVTSSVPLEHLNDSKRKILDERTAELINVLQIMSTL"
2771..2776
/citation=[2]
/evidence=experimental
2794..2799
/citation=[2]
/evidence=experimental
2874..3377
/gene="qcrA"
2874..3377
/gene="qcrA"
/function="menaquinone oxidase"
/standard_name="Rieske protein"
/note="26.2% of identical aminoacids with the
cyanobacterium Nostoc PCC 7906 Rieske iron-sulfur protein"
/citation=[2]
/codon_start=1
/evidence=experimental
/transl_table=11
/product="Rieske iron-sulfur protein"
/protein_id="AAB38435.1"
/db_xref="GI:1146227"
/translation="MGCKHDSRRQELNYTLNGVGRMASMLMPWRFPALDPLVAST
GKQDMQVVSVDLTKBPQDFPRKINQVDAWYSEBSRSAMVVRKNGEITVALSPICH
LGCTVMNNSDPKNNPKFPCCHYGLYKQSTNVGTPPLALDHYEQEVKQGLYLIGK
AKPKGEG"
3379..4053
/gene="qcrB"
3379..4053
/gene="qcrB"
/function="cytochrome reductase"

gene
CDS

gene
CDS

standard_name="cytochrome b6"
/note="42.3% of identical aminoacids with Chlorella
proteocoides chloroplast cytochrome b6"
/citation=[2]
/codon_start=1
/evidence=experimental
/transl_table=11
/product="cytochrome b"
/protein_id="AAB38436.1"
/db_xref="GI:1146228"
/translation="MLNKIYDWDERLDITPMWRDIDADHEPEHVNPAHFSAFYCF
GGLPPTLVQVSGMKPLTWYVPDIQNAEVSYYLONEVAFCOIVRGHMWASLVI
VMFHLTRLPFGQAAYKKPRRLNIYGVLLFFMNLGIGFTGYLLPMDMALPATKXGL
QIAATPLTIGVOVTLLAGHPDIVAQTLTRFAIHFFILPALFGILMAHPIINIKQ
GISGPL"

4088..4855
/gene="qcrC"
4088..4855
/gene="qcrC"
/note="29.1% of identical aminoacids with the Spinacia
oleracea chloroplast cytochrome b"
/citation=[2]
/codon_start=1
/evidence=experimental
/transl_table=11
/product="cytochrome b"
/protein_id="AAB38437.1"
/db_xref="GI:1146229"
/translation="MHREGKMKFVGDSRIIPAEEKNIPXDYSYEGKTEAFMPENFLIK
EMMGAVNLIGFLVITYTHQPPLERMDDPDNDGYILPDWYPLFIYOLLXEVAASF
TVGGMIPGIALGALLAPPLDRGTERRPKRPVANVGMILLISAAPVLTNOSVATH
DMAKEBGGKTTEKDIDITNAEGYKVFKEQGCSCHGDNLQGAAGSLVDGSKDE
IKKIYAEBKGKMPAGVFPKSNDKOLEBLAKFISETTA"

4898..4922
/standard_name="transcription terminator"
/note="putative"
/citation=[2]

gene
CDS

5002..5559
/gene="ypjA"
5002..5559
/gene="ypjA"
/function="hypothetical"
/note="putative"
/citation=[2]
/codon_start=1
/transl_table=11
/protein_id="AAB38438.1"
/db_xref="GI:1146230"
/translation="MLILVLAINFLGVYGYWTLPOLLETPARFLIFDPDSPATPF
FLVYLAIAPFKMNAPLSLEALALVTLVKYGAAVAMNFVLVALVTGDLPEWGMLIASHF
FMVAOGVSYEPFRSFWHALIAVAWTLHANDVIDYLFDMWPQYSMLSMDYTELGTFP
WTSISLAIVLAVLYSSKKQTGLELM"

5628..6422
/gene="ypjB"
5628..6422
/gene="ypjB"
/function="hypothetical"
/note="putative"
/citation=[2]
/codon_start=1
/transl_table=11
/protein_id="AAB38439.1"
/db_xref="GI:1146231"
/translation="MKRKLTICLLALIFPNGNAKAERGSLEELINDLSDTPQMTQQ
ACYBERLOUTLEYPKTKISAEEKODPMLTGAOIROITLYNDWVRSLKQDTSDTOKR
LBAAPQFRMLDAVNDSPDLMSLEKPIMARFTBLKDVQKNSTSHKEMNFSIL
YDLIPLSLTIYVSDQLETVNGKHIDIVEQEDEFQQTSTFKLERUSLQHDIAKNVFDKY
EEBDADPSLIHWIIITGSIITRALTYVGRRKKYKRNKLKRDRYP"

6434..6447
/standard_name="transcription terminator"
/note="putative"
/citation=[2]
complement(6464..7111)

terminator

gene

CDS	/gene="ypjC"	/complement(6464..7111)
	/gene="ypjC"	/function="hypothetical"
	/note="putative"	
Query Match	100.0%; Score 494; DB 1; Length 23775;	
Beeb Local Similarity	100.0%; Pred. No. 3,6e-123; Indels 0; Gaps 0;	
Matches 494; Conservative	0; Mismatches 0; Indels 0; Gaps 0;	
OY	1	GCTAATGTGGTGTACGAAGCCCGTTGATTGGTAATCTTCATTGGGCAATCGCTG 60
DB	13043	GCTAACTGTGGTGTACGAAGCCCGTTGATTGGTAATCTTCATTGGGCAATCGCTG 13102
OY	61	CGAACTGCACCTTATTTATAAATAGATAGAATTGCAAGAGTCCTGATCCAAAAA 120
DB	13103	CGAACTGCACCTTATTTATAAATAGATAGAATTGCAAGAGTCCTGATCCAAAAA 13162
OY	121	GGACTGGGACAAGGGGATGAAACTGCCCGAACCTTTAGAAAGTGAAATCCTTCGTTG 180
DB	13163	GGACTGGGACAAGGGGATGAAACTGCCCGAACCTTTAGAAAGTGAAATCCTTCGTTG 13222
OY	181	TAA CGAAGGTTTTTTGGCTTGGCAAGAAAACGGCAGATCATCTCTTAACATGAGG 240
DB	13223	TAA CGAAGGTTTTTTGGCTTGGCAAGAAAACGGCAGATCATCTCTTAACATGAGG 13282
OY	241	AGGAGAAAACATGAAAACAAAACCTGAGTTTTCTTAAAAATGAAAGAGTCGAAAGAACCAT 300
DB	13283	AGGAGAAAACATGAAAACAAAACCTGAGTTTTCTTAAAAATGAAAGAGTCGAAAGAACCAT 13342
OY	301	TGTCATCTGACCGGCTTATGATATCCGAGCTAAACTTGCTGAAACAAGCGGAGTTGA 360
DB	13343	TGTCATCTGACCGGCTTATGATATCCGAGCTAAACTTGCTGAAACAAGCGGAGTTGA 13402
OY	361	CATGATTTTAGTCGGTGATTCACCTTGGAAATGTCGTCTCGGCTTGATTCATCTGCGG 420
DB	13403	CATGATTTTAGTCGGTGATTCACCTTGGAAATGTCGTCTCGGCTTGATTCATCTGCGG 13462
OY	421	TGTGACAGTTGCGGACATGATTCATCTATACAAAACCGCTTAAAAAGGGTGCCTCCGAATTC 480
DB	13463	TGTGACAGTTGCGGACATGATTCATCTATACAAAACCGCTTAAAAAGGGTGCCTCCGAATTC 13522
OY	481	CTTATATTGTGACAG 494
DB	13523	CTTATATTGTGACAG 13536
RESULT 3		
BSUB0012/c	201375 bp	DNA linear BCT 07-JUL-2003
LOCUS	BSUB0012	Bacillus subtilis complete genome (section 12 of 21) : from 2207806
DEFINITION		to 2409180.
ACCESSION	299115 AL0009126	
VERSION	299115.2	GI:32468778
KEYWORDS		
SOURCE ORGANISM		Bacillus subtilis subsp. subtilis str. 168
REFERENCE		Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
AUTHORS		1 (bases 1 to 201375) Kunst, F., Ogasawara, N., Moszer, I., Albertini, A.M., Alloni, G., Azavedo, V., Bertero, M.G., Bessieres, P., Bolotin, A., Borcherdt, S., Borries, R., Boutsier, L., Brans, A., Braun, M., Bricogne, S.C., Bron, S., Broillet, S., Bruschi, C.V., Caldwell, B., Capuano, V., Carter, N.M., Choi, S.K., Codani, J.J., Conneron, I.F., Cummings, N.J., Daniel, R.A., Denizot, F., Devine, K.M., Dusterhoft, A., Ehrlich, S.D., Emmeron, P.T., Entian, K.D., Errington, J., Fabret, C., Ferrieri, E., Foulger, D., Fritz, C., Fujita, M., Fujita, Y., Fuma, S., Galizzi, A., Galleron, D., Ghim, S.Y., Glaeser, P., Goffeau, A., Golightly, E.J., Grandi, G., Halbepré, Guy, B.J., Haga, K., Hahsch, U., Harwood, C.R., Hénaut, A., Hilbert, H., Holtsappel, S., Hosono, S., Huilo, M.F., Klaer, M., Jones, L., Joris, B., Karamata, D., Kasahara, Y., Kleary-Blochard, M., Klein, C., Kobayashi, Y., Koetter, P., Konigstein, G., Krogh, S., Kumano, M., Kurita, K., Lapidus, A.,

gene complement (3636. .3887)
/gene="yopB"
/locus_tag="BSU20920"
complement (3636. .3887)
/gene="yopB"
/locus_tag="BSU20920"
/function="unknown"
/codon_start=1
/transl_table=1
/protein_id="CAB14010.1"
/db_xref="GI:2634512"
/db_xref="UniProt/TREMBL:O31933"
/translation="MIGLATFLTWLGVGLTGKIFVQVYVDEEKLMDKETAAG
MERLALSLFKNKLNVIAFEMLIGLPLAMKITLPRG"
complement (4474. .4869)
/gene="yopD"
/locus_tag="BSU20930"
complement (4474. .4869)
/gene="yopD"
/locus_tag="BSU20930"
/function="unknown"
/locus_tag="BSU20930"
/function="unknown"
/note="similar to unknown proteins from B. subtilis"
/codon_start=1
/transl_table=1
/protein_id="CAB14011.1"
/db_xref="GI:2634513"
/db_xref="UniProt/TREMBL:O31934"
/translation="MSLNQKALNSVYVTLAFSLFGLIFGLVYVSGFWALVTI
ALIAFVLLITVLPAAPIQVWLRRRRRFSLINFLIYAVAFSAVLFVWFVDPYDNL
TFRSPFYINISVIAPIYFWDSIFLRN"
Query Match 100.0%; Score 494; DB 1; Length 201375;
Best Local Similarity 100.0%; Pred. No. 3.3e-123;
Matches 494; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
CDS
gene
CDS
1 GCTAATGCTGGTGCATCAAGCCCGTGGTATTTGGTACTTCCATTTGGGAGATTCGCTG 60
Db 146314 GCTAATGCTGGTGCATCAAGCCCGTGGTATTTGGTACTTCCATTTGGGAGATTCGCTG 146255
Qy 61 GGAAGTGCACCTATATTAATAATAGATAGACATTCGACAGTCTGCTTGAATCCAAAAA 120
Db 146254 GGAAGTGCACCTATATTAATAATAGATAGACATTCGACAGTCTGCTTGAATCCAAAAA 146195
Qy 121 GGAAGTGCACAGAGGATGAATCTGCCGAATTTAGAAAGTGAAGATCTTCTGTTG 180
Db 146194 GGAAGTGCACAGAGGATGAATCTGCCGAATTTAGAAAGTGAAGATCTTCTGTTG 146135
Qy 181 TAAAGGAAAGTCTTGGCTTGCAGAGAAACGCAATCATCTCTCTTAAACATGAGG 240
Db 146134 TAAAGGAAAGTCTTGGCTTGCAGAGAAACGCAATCATCTCTCTTAAACATGAGG 146075
Qy 241 AGGAGAAACATGAACAAACGATTTCTAATAAATGAAGAGTCTGAAGAACCAT 300
Db 146074 AGGAGAAACATGAACAAACGATTTCTAATAAATGAAGAGTCTGAAGAACCAT 146015
Qy 301 TGTCAATGCTGACCGCTTATGATTAATCCGACAGTAACTTGTGAAACAGCGGAGTTGA 360
Db 146014 TGTCAATGCTGACCGCTTATGATTAATCCGACAGTAACTTGTGAAACAGCGGAGTTGA 145955
Qy 361 CATGATTTTATGTCGGTATTCATCTTGAATGCTGCTCGGCTTGAATCACTGTGG 420
Db 145954 CATGATTTTATGTCGGTATTCATCTTGAATGCTGCTCGGCTTGAATCACTGTGG 145895
Qy 421 TGTGACAGTGGGACATGATCATCATCAAAAGCGTTAAAGGGGTGCGCGAATAC 480
Db 145894 TGTGACAGTGGGACATGATCATCATCAAAAGCGTTAAAGGGGTGCGCGAATAC 145835
Qy 481 CTTTATTTGTGACAG 494
Db 145834 CTTTATTTGTGACAG 145821
RESULT 4

AX100815
LOCUS AX100815 2363 bp DNA linear PAT 10-APR-2001
DEFINITION Sequence 59 from Patent WO0121772.
ACCESSION AX100815
VERSION AX100815.1 GI:13619747
KEYWORDS
SOURCE
ORGANISM Bacillus subtilis
Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
REFERENCE
1 Yocum,R.R., Patterson,T.A., Hermann,T. and Pero,J.G.
TITLE Methods and microorganisms for production of panto-compounds
JOURNAL Patent: WO 0121772-A 59 29-MAR-2001;
OMNIGENE BIOPRODUCTS, INC. (US)
FEATURES
source
CDS
CDS
1 1077. .1937
/note="unnamed protein product"
/codon_start=1
/transl_table=1
/protein_id="CAC36486.1"
/db_xref="GI:13619749"
/translation="MROITDISQKEAIKQYHSRKSIGFVPMGFILHGLTLADRA
RQENDAVIMSIPVNPAPQFQEDDEAVPRIDERDALABNADVILTPPAHMPRE
KNVTHIVERETDVLGSRSGHEDGVAIVLTKLNLVKTPTAFGLDAQOAVAVDL
ISDFMDIELVAVDTVEEDGLAKSSNNVLTLEERKAPLXRALQTSLELVQGER
DPEAVIKAKDIIETTSITIDVELYSYPELIPVNEIAGKMLAVAAVAFSRILDMI
IIDIREMER"
1939. .2322
/note="unnamed protein product"
/codon_start=1
/transl_table=1
/protein_id="CAC36487.1"
/db_xref="GI:13619750"
/translation="MYRTMSGKLHRAVTEANLVYGSITIDEDLIDAVGMLPNEKY
QIVNNNGARLEITLIPGKSGSVTCNAGAAARLVGRDKVITIIISYKMSDQEAASHE
PKVAVLNDONKIEQMLGNEPARTL"
ORIGIN
Query Match 98.2%; Score 485; DB 6; Length 2363;
Best Local Similarity 100.0%; Pred. No. 1.1e-120;
Matches 485; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
CDS
CDS
10 TTGGTACAAGCCCGTGGTATTTGGTATTAATCTTCATTTGGGAGATTCGCTGCAATCTGCA 69
Db 1 TTGGTACAAGCCCGTGGTATTTGGTATTAATCTTCATTTGGGAGATTCGCTGCAATCTGCA 60
Qy 70 CCTATTATTAATAATAGATAGACATTCGACAGTCTGCTTGAATCCAAAAAGACTGGGA 129
Db 61 CCTATTATTAATAATAGATAGACATTCGACAGTCTGCTTGAATCCAAAAAGACTGGGA 120
Qy 130 CAGAGGATGAACCTCGCGCAACTTTGAAGTGAAGAACTCTCTGTTGAACGGAAG 189
Db 121 CAGAGGATGAACCTCGCGCAACTTTGAAGTGAAGAACTCTCTGTTGAACGGAAG 180
Qy 190 GTTTTGGCTTGCAGAGAAACGCGAGATCATCTCTCTTAAACATGAGAGAGAGAAA 249
Db 181 GTTTTGGCTTGCAGAGAAACGCGAGATCATCTCTCTTAAACATGAGAGAGAGAAA 240

LOCUS AX572874 6725 bp DNA linear PAT 29-NOV-2002
DEFINITION Sequence 22 from Patent WO02057476.
ACCESSION AX572874
VERSION AX572874.1 GI:26004960
KEYWORDS
SOURCE
ORGANISM
REFERENCE
1 Hermann, T., Patterson, T.A., Pero, J.G., Yocum, R.R., Baldeusius, K.U. and Beck, C. Methods and microorganisms for the production of 3-(2-hydroxy-3-methyl-butylamino)-propionic acid (hmbpa) Patent: WO 02057476-A 22 25-JUL-2002;
JOURNAL Omnigene Bioproducts (US)
FEATURES
LOCATION/Qualifiers
SOURCE
1. .6725
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="vector-pam636"

Db 250 CATGAAACAAACATGATTTCTTAAATGAGAGCTGAGAACCGATTGTCATGCT
241 CATGAAACAAACATGATTTCTTAAATGAGAGCTGAGAACCGATTGTCATGCT
Qy 310 GACCGCTTATGATTAATCCGGAGCTAACTTGCTGAACAGCGGAGTGAATGATTT
Db 301 GACCGCTTATGATTAATCCGGAGCTAACTTGCTGAACAGCGGAGTGAATGATTT
Qy 370 AGTCGCTATTCATCTTGAAATGCTGCTCGGCTTATTCACCTGTCGGTGTGACAGT
Db 361 AGTCGCTATTCATCTTGAAATGCTGCTCGGCTTATTCACCTGTCGGTGTGACAGT
Qy 430 TCGCGACATGATCCATCAACAAAGCCGTTAAAGGGGTGCGCGAATACCTTATTTGT
Db 421 TCGCGACATGATCCATCAACAAAGCCGTTAAAGGGGTGCGCGAATACCTTATTTGT
Qy 490 GACAG 494
Db 481 GACAG 485

RESULT 5
AX100850/c 7381 bp DNA linear PAT 10-APR-2001
LOCUS AX100850
DEFINITION Sequence 94 from Patent WO0121772.
ACCESSION AX100850
VERSION AX100850.1 GI:13619778
KEYWORDS
SOURCE
ORGANISM
REFERENCE
1 Yocum, R.R., Patterson, T.A., Hermann, T. and Pero, J.G. Methods and microorganisms for production of panto-compounds Patent: WO 0121772-A 94 29-MAR-2001;
JOURNAL OMNIGENE BIOPRODUCTS, INC. (US)
FEATURES
LOCATION/Qualifiers
SOURCE
1. .7381
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="Description of Artificial Sequence: recombinant PAN006 plasmid"

ORIGIN

Query Match 52.2%; Score 258; DB 6; Length 7381;
Best Local Similarity 100.0%; Pred. No. 4.5e-59; Indels 0; Gaps 0;
Matches 258; Conservative 0; Mismatches 0;

Qy 237 GAGGAGGAGAAACATGAAACAAACATGATTTCTTAAATGAGAGCTGAGAAC 296
Db 3652 GAGGAGGAGAAACATGAAACAAACATGATTTCTTAAATGAGAGCTGAGAAC 3593
Qy 297 CGATTGTCATGCTGACCGCTTATGATTAATCCGGAGCTAAACTTGCTGAACAGCGGAG 356
Db 3592 CGATTGTCATGCTGACCGCTTATGATTAATCCGGAGCTAAACTTGCTGAACAGCGGAG 3533
Qy 357 TTGACATGATTTTATGTCGCTGATTCATCTTGGAATGCTGCTCGGCTTATTAATG 416
Db 3532 TTGACATGATTTTATGTCGCTGATTCATCTTGGAATGCTGCTCGGCTTATTAATG 3473
Qy 417 TCGGTGTGACAGTTCGCGACATGATCCATCAACAAAGCCGTTAAAGGGGTGCGCGA 476
Db 3472 TCGGTGTGACAGTTCGCGACATGATCCATCAACAAAGCCGTTAAAGGGGTGCGCGA 3413
Qy 477 ATACCTTATTTGTGACAG 494
Db 3412 ATACCTTATTTGTGACAG 3395

RESULT 6
AX572874

LOCUS AX572874 6725 bp DNA linear PAT 29-NOV-2002
DEFINITION Sequence 22 from Patent WO02057476.
ACCESSION AX572874
VERSION AX572874.1 GI:26004960
KEYWORDS
SOURCE
ORGANISM
REFERENCE
1 Hermann, T., Patterson, T.A., Pero, J.G., Yocum, R.R., Baldeusius, K.U. and Beck, C. Methods and microorganisms for the production of 3-(2-hydroxy-3-methyl-butylamino)-propionic acid (hmbpa) Patent: WO 02057476-A 22 25-JUL-2002;
JOURNAL Omnigene Bioproducts (US)
FEATURES
LOCATION/Qualifiers
SOURCE
1. .6725
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="vector-pam636"

ORIGIN

Query Match 52.0%; Score 257; DB 6; Length 6725;
Best Local Similarity 100.0%; Pred. No. 8.5e-59; Indels 0; Gaps 0;
Matches 257; Conservative 0; Mismatches 0;

Qy 238 AGGAGGAGAAACATGAAACAAACATGATTTCTTAAATGAGAGCTGAGAAC 297
Db 325 AGGAGGAGAAACATGAAACAAACATGATTTCTTAAATGAGAGCTGAGAAC 384
Qy 298 GATTGTCATGCTGACCGCTTATGATTAATCCGGAGCTAAACTTGCTGAACAGCGGAGT 357
Db 385 GATTGTCATGCTGACCGCTTATGATTAATCCGGAGCTAAACTTGCTGAACAGCGGAGT 444
Qy 358 TGACATGATTTTATGTCGCTGATTCATCTTGGAATGCTGCTCGGCTTATTAATG 417
Db 445 TGACATGATTTTATGTCGCTGATTCATCTTGGAATGCTGCTCGGCTTATTAATG 504
Qy 418 CGGTGTGACAGTTCGCGACATGATCCATCAACAAAGCCGTTAAAGGGGTGCGCGA 477
Db 505 CGGTGTGACAGTTCGCGACATGATCCATCAACAAAGCCGTTAAAGGGGTGCGCGA 564
Qy 478 TACCTTATTTGTGACAG 494
Db 565 TACCTTATTTGTGACAG 581

RESULT 7
AX573226 6725 bp DNA linear PAT 29-NOV-2002
LOCUS AX573226
DEFINITION Sequence 22 from Patent WO02057474.
ACCESSION AX573226
VERSION AX573226.1 GI:26005139
KEYWORDS
SOURCE
ORGANISM
REFERENCE
1 Hermann, T., Patterson, T.A., Pero, J.G., Yocum, R.R., Baldeusius, K.U. and Beck, C. Processes for enhanced production of pantothenate Patent: WO 02057474-A 22 25-JUL-2002;
JOURNAL OMNIGENE BIOPRODUCTS, INC. (US)
FEATURES
LOCATION/Qualifiers
SOURCE
1. .6725
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="vector-pam636"

ORIGIN

Query Match 52.0%; Score 257; DB 6; Length 6725;

Query	Best Local Similarity	100.0%	Pred. No. 8.5e-59;	Matches 257;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0;
QY	238	AGGAGGAGAAAA	CATGAAAA	CAAAACTGGATTTTCTTAAAAATGAAGAGCTGAAGAACC	297							
Db	325	AGGAGGAGAAAA	CAATGAAAA	CAAAACTGSAATTTCTTAAAAATGAAGAGCTGAAGAACC	384							
QY	298	GATTGTCAATGCTGA	CCGCTTATGATTTATCCGGACGCTAAACTTGTCTGAACAAGGGGGAGT	357								
Db	385	GATTGTCAATGCTGA	CCGCTTATGATTTATCCGGACGCTAAACTTGTCTGAACAAGGGGGAGT	444								
QY	358	TGACATGATTTTAACT	CGGTGATTTCACTTGAATGGTGGTCTCGGCTTGAATCAACTGT	417								
Db	445	TGACATGATTTTAACT	CGGTGATTTCACTTGAATGGTGGTCTCGGCTTGAATCAACTGT	504								
QY	418	CGGTGTGACAGTTGCGG	ACATGATTCATATACAAAAGCCGTTAAAAAGGGGTGCGCGAA	477								
Db	505	CGGTGTGACAGTTGCGG	ACATGATTCATATACAAAAGCCGTTAAAAAGGGGTGCGCGAA	564								
QY	478	TACCTTTATTTGTGACAG	494									
Db	565	TACCTTTATTTGTGACAG	581									
RESULT 8	AX100849/c	8503 bp	DNA	linear	PAT 10-APR-2001							
LOCUS	AX100849	Sequence 93 from Patent WO0121772.										
DEFINITION	AX100849											
ACCESSION	AX100849											
VERSION	AX100849.1	GI.13619777										
KEYWORDS												
SOURCE												
ORGANISM												
REFERENCE	1											
AUTHORS	Yocum,R.R., Paterson,T.A., Hermann,T. and Pero,J.G.											
TITLE	Methods and microorganisms for production of panto-compounds											
JOURNAL	Patent: WO 0121772-A 93 29-MAR-2001;											
FEATURES	OMNIGENE BIOPRODUCTS, INC. (US)											
source	location/Qualifiers											
	1..8503											
	/organism="synthetic construct"											
	/mol_type="unassigned DNA"											
	/db_xref="taxon:32630"											
	/note="Recombinant pAN004 plasmid"											
ORIGIN												
Query Match	52.0%;	Score 257;	DB 6;	Length 8503;								
Best Local Similarity	100.0%;	Pred. No. 8.4e-59;										
Matches 257;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0;				
QY	238	AGGAGGAGAAAA	CATGAAAA	CAAAACTGGATTTTCTTAAAAATGAAGAGCTGAAGAACC	297							
Db	3656	AGGAGGAGAAAA	CAATGAAAA	CAAAACTGSAATTTCTTAAAAATGAAGAGCTGAAGAACC	359							
QY	298	GATTGTCAATGCTGA	CCGCTTATGATTTATCCGGACGCTAAACTTGTCTGAACAAGGGGGAGT	357								
Db	3596	GATTGTCAATGCTGA	CCGCTTATGATTTATCCGGACGCTAAACTTGTCTGAACAAGGGGGAGT	353								
QY	358	TGACATGATTTTAACT	CGGTGATTTCACTTGAATGGTGGTCTCGGCTTGAATCAACTGT	417								
Db	3536	TGACATGATTTTAACT	CGGTGATTTCACTTGAATGGTGGTCTCGGCTTGAATCAACTGT	347								
QY	418	CGGTGTGACAGTTGCGG	ACATGATTCATATACAAAAGCCGTTAAAAAGGGGTGCGCGAA	477								
Db	3476	CGGTGTGACAGTTGCGG										

AX100832/c	AX100832	10801 bp	DNA	linear	PAT 10-APR-2001
LOCUS	Sequence 76 from Patent WO0121772.				
DEFINITION	AX100832				
ACCESSION	AX100832.1	GI:13619763			
VERSION					
KEYWORDS					
SOURCE					
ORGANISM					
REFERENCE					
AUTHORS	1				
TITLE	Yocum,R.R., Patterson,T.A., Hermann,T. and Pero,J.G.				
JOURNAL	Methods and microorganisms for production of panto-compounds				
	Patent: WO 0121772-A 76 29-MAR-2001;				
	OMNIGENE BIOPRODUCTS, INC. (US)				
FEATURES					
source	location/Qualifiers				
	1..10801				
	/organism="synthetic construct"				
	/mol_type="unassigned DNA"				
	/db_xref="taxon:12630"				
	/note="Recombinant PAN240 plasmid"				
ORIGIN					
Query Match	52.0%;	Score 257;	DB 6;	Length 10801;	
Best Local Similarity	100.0%;	Pred. No. 8.3e-59;			
Matches 257;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;	
QY	238	AGGAGGAGAAAACATGAAAAACAAAACCTGATTTCTTAAATGGAAGAGTGTGAAGAACC	297		
Db	3656	AGGAGGAGAAAACATGAAAAACAAAACCTGATTTCTTAAATGGAAGAGTGTGAAGAACC	3597		
QY	298	GATTGTCACTGTGACCGCTTATGATTCGCGACAGCTAACTTCCTGTAACAAGCGGAGT	357		
Db	3586	GATTGTCACTGTGACCGCTTATGATTCGCGACAGCTAACTTCCTGTAACAAGCGGAGT	3533		
QY	358	TGACATGATTTTAACTCGGTGATTCACCTTGAATGCTGCTCGGCTTGAATCACTGT	417		
Db	3536	TGACATGATTTTAACTCGGTGATTCACCTTGAATGCTGCTCGGCTTGAATCACTGT	3477		
QY	418	CGGTGTGACAGTTCGGACATGATTCATCAACAAAACCGCTTAAAGGGGTGGCGCGAA	477		
Db	3476	CGGTGTGACAGTTCGGACATGATTCATCAACAAAACCGCTTAAAGGGGTGGCGCGAA	3417		
QY	478	TACCTTTATGTGACAG 494			
Db	3416	TACCTTTATGTGACAG 3400			
RESULT 10					
AX100779	AX100779	831 bp	DNA	linear	PAT 10-APR-2001
LOCUS	Sequence 23 from Patent WO0121772.				
DEFINITION	AX100779				
ACCESSION	AX100779.1	GI:13619711			
VERSION					
KEYWORDS					
SOURCE					
ORGANISM					
REFERENCE					
AUTHORS	1				
TITLE	Yocum,R.R., Patterson,T.A., Hermann,T. and Pero,J.G.				
JOURNAL	Methods and microorganisms for production of panto-compounds				
	Patent: WO 0121772-A 23 29-MAR-2001;				
	OMNIGENE BIOPRODUCTS, INC. (US)				
FEATURES					
source	location/Qualifiers				
	1..831				
	/organism="Bacillus subtilis"				
	/mol_type="unassigned DNA"				
	/db_xref="taxon:1423"				
	1..>831				
	/note="unamed protein product"				
	/codon_start=1				
	/transl_table=11				
	/protein_id="CAC36477.1"				
	/db_xref="GI:13619712"				

/translation="MKTKLDFLKKKSEEPILVLTAYDYPAAKLAEOAGVMLIVGDS
LGMVTLGDSITVGTVAIDMIHTKAVKGAAPRTIIVDMPEMSVHLSKEOTLKRAAI
VQESGADPLKLEGGSGVPEIRALITLGGIPVSHLGLTPGVGLGKYGKRGQA
KKLIEDSTKCEAGMMVLPCVPAELTAKIAETLSIPVIGIGAVKADGVLYHDI
IGHVERTPKFVKQYTRIDETIETASISGVODVHRHRAPEQKHSFQNMQTVLDGLYGG
K"

ORIGIN

Query Match 49.4%; Score 244; DB 6; Length 811;
Best Local Similarity 100.0%; Pred. No. 3.2e-55;

Matches 244; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 251 ATGAAACAAAACAGTATTTCTTAAATGAAGAGTCTGAAGAACGATTTGTCATGCTG 310
Db 1 ATGAAACAAAACAGTATTTCTTAAATGAAGAGTCTGAAGAACGATTTGTCATGCTG 60
Qy 311 ACCGCTTATGATTAATCCGCGACCTAAACTTGTGAACAACGGGAGTTGACATGATTTTA 370
Db 61 ACCGCTTATGATTAATCCGCGACCTAAACTTGTGAACAACGGGAGTTGACATGATTTTA 120
Qy 371 GTCGATGATTCATCTGGAAATGCTGCTCGGCTTGAATTCATCTGCTGCTGACACTT 430
Db 121 GTCGATGATTCATCTGGAAATGCTGCTCGGCTTGAATTCATCTGCTGCTGACACTT 180
Qy 431 GCGGACATGATTCATCATACAAAGCCGTTAAAGGGGCGCCGATACCTTTATTTG 490
Db 181 GCGGACATGATTCATCATACAAAGCCGTTAAAGGGGCGCCGATACCTTTATTTG 240
Qy 491 ACAG 494
Db 241 ACAG 244

RESULT 11
AE017333_23/c
WPCOMMENT

Sequence split into 43 fragments LOCUS AE017333 Accession AE017333

Fragment Name	Begin	End
AE017333_00	1	110000
AE017333_01	100001	210000
AE017333_02	200001	310000
AE017333_03	300001	410000
AE017333_04	400001	510000
AE017333_05	500001	610000
AE017333_06	600001	710000
AE017333_07	700001	810000
AE017333_08	800001	910000
AE017333_09	900001	1010000
AE017333_10	1000001	1110000
AE017333_11	1100001	1210000
AE017333_12	1200001	1310000
AE017333_13	1300001	1410000
AE017333_14	1400001	1510000
AE017333_15	1500001	1610000
AE017333_16	1600001	1710000
AE017333_17	1700001	1810000
AE017333_18	1800001	1910000
AE017333_19	1900001	2010000
AE017333_20	2000001	2110000
AE017333_21	2100001	2210000
AE017333_22	2200001	2310000
AE017333_23	2300001	2410000
AE017333_24	2400001	2510000
AE017333_25	2500001	2610000
AE017333_26	2600001	2710000
AE017333_27	2700001	2810000
AE017333_28	2800001	2910000
AE017333_29	2900001	3010000
AE017333_30	3000001	3110000
AE017333_31	3100001	3210000
AE017333_32	3200001	3310000
AE017333_33	3300001	3410000
AE017333_34	3400001	3510000

AE017333_35 3500001 3610000
AE017333_36 3600001 3710000
AE017333_37 3700001 3810000
AE017333_38 3800001 3910000
AE017333_39 3900001 4010000
AE017333_40 4000001 4110000
AE017333_41 4100001 4210000
AE017333_42 4200001 4222645
Continuation (24 of 43) of AE017333 from base 2300001 (AE017333 Bacillus licheniformis DK

Query Match 46.4%; Score 229; DB 1; Length 110000;
Best Local Similarity 73.3%; Pred. No. 3e-51;

Matches 366; Conservative 0; Mismatches 115; Indels 18; Gaps 5;

Qy 9 GTTGATCAAGCCCGTGAATTTGGTATTAATTCATTCATGGGCAATACGCTGGAAGTGC 68
Db 25528 GTTGAGACGATCCATGCGTTTGGTATTAATTCATTCATGGGCAATACGCTGGAAGTGC 25469
Qy 69 ACTTATTAATTAATATAGATGACATTCGACAGTCTGCTTGAATCCAAAAAGACTGCG 128
Db 25468 ACTTATTAATTAATATAGATGACATTCGACAGTCTGCTTGAATCCAAAAAGACTGCG 25411
Qy 129 ACAGAGGATGAACCTGCCGAACCTTAGAAGTGAAGATCCTTCTGTTAGCGAA 188
Db 25410 ACAGAGGATGAATCTTGCCGAAC-ACGAGCATAGCAAAATCCTTCTGTTAGCGAA 25354
Qy 189 GGTATTTTGGCTTGCAGAAAGAA-----ACGCGATGATCATCTCTTAACA 225
Db 25353 GGTATTTTGGCTTGCAGAAAGATGCAATGCAAGCAGGACATCATCTCTGTAAGAA 25294
Qy 236 TGAGAGAGAGAAAAATGAAACAAACTGATTTTCTAAAAATGAAGAGTGAAGAA 295
Db 25293 AGAGAGAGAGAAAAATGAAACAAACTGATTTTCTAAAAATGAAGAGAGTGAAGAA 25234
Qy 296 CCGATTGTCACTGTCAGCCGCTTATGATTAATTCGCGCAGCTAAACTTGTGAACAACGCGGA 355
Db 25233 CCGATTGTCACTGTCAGCTGCTTATGATTAATTCGCGCAGCTAAACTTGTGTGAACAACGCGCG 25174
Qy 356 GTTGACATGATTTTATGTCGCTGATTCATCTGGAATGCTGCTCGGCTTGAATTAAGT 415
Db 25173 GTTGACATGATTTTATGTCGCTGATTCATCTGGAATGCTGCTCGGCTTGAATTAAGT 25114
Qy 416 GTCGATGATGACAGTTCGCGACATGATTCATCAAAAGCCGTTAAAGGGTGCAGCG 475
Db 25113 GTCGATGATGACAGTTCGCGACATGATTCATCAAAAGCCGTTAAAGGGTGCAGCG 25054
Qy 476 AATACCTTAATGTCAGAG 494
Db 25053 AATACCTTAATGTCAGAG 25035

RESULT 12
CP000002_23/c
WPCOMMENT

Sequence split into 43 fragments LOCUS CP000002 Accession CP000002

Fragment Name	Begin	End
CP000002_00	1	110000
CP000002_01	100001	210000
CP000002_02	200001	310000
CP000002_03	300001	410000
CP000002_04	400001	510000
CP000002_05	500001	610000
CP000002_06	600001	710000
CP000002_07	700001	810000
CP000002_08	800001	910000
CP000002_09	900001	1010000
CP000002_10	1000001	1110000
CP000002_11	1100001	1210000
CP000002_12	1200001	1310000
CP000002_13	1300001	1410000
CP000002_14	1400001	1510000
CP000002_15	1500001	1610000
CP000002_16	1600001	1710000

CP000002_17 1700001 1810000
CP000002_18 1800001 1910000
CP000002_19 1900001 2010000
CP000002_20 2000001 2110000
CP000002_21 2100001 2210000
CP000002_22 2200001 2310000
CP000002_23 2300001 2410000
CP000002_24 2400001 2510000
CP000002_25 2500001 2610000
CP000002_26 2600001 2710000
CP000002_27 2700001 2810000
CP000002_28 2800001 2910000
CP000002_29 2900001 3010000
CP000002_30 3000001 3110000
CP000002_31 3100001 3210000
CP000002_32 3200001 3310000
CP000002_33 3300001 3410000
CP000002_34 3400001 3510000
CP000002_35 3500001 3610000
CP000002_36 3600001 3710000
CP000002_37 3700001 3810000
CP000002_38 3800001 3910000
CP000002_39 3900001 4010000
CP000002_40 4000001 4110000
CP000002_41 4100001 4210000
CP000002_42 4200001 4222336
Continuation (24 of 43) of CP000002 from base 2300001 (CP000002 Bacillus licheniformis D

Query Match 46.4%; Score 229; DB 1; Length 110000;

Best Local Similarity 73.3%; Pred. No. 3e-51; Mismatches 115; Indels 18; Gaps 5;

Matches 366; Conservative 0;

9 GTTGTACAGCCCTGATTTTGTATATCTTCATTGGGCGATTCGCTGCGAATCTG 68
24670 GTTGTACAGATCCAGCGTTTGTATATCTTCATTGGGCGATTCGCTGCGAATCTG 24611
69 ACCATATTAATAATAGATAGACATTGACAGCTGCTGCTGATCCAAAAAGAGCTGG 128
24610 ACCGTAAAGCATGCACTT-ATAAAAAACAGTCGCTGATCC-AAAAAGAGCTGG 24553
129 ACAGAGGATGAACTCCGCCAATTTAGAAAGTGAAGATCTTCGTTGTAACGGA 188
24552 ACAGAGGATGATTTCTTCCGAC-ACGAGCATACGAATCTTCTCTT-CCGAA 24496
189 GCTTTTGGCTTGCAAGAA-----ACGCGATCATCTCCTTAAACA 235
24495 GCTTTTGGCTTGCAAGAAAGATGCAAGAGCGGCTCATCTCTGAAAA 24436
236 TGAGAGAGAAACATGAAACAAACTGATTTTCTAATAATGAAGAGTCTGAAGA 295
24435 AGAGAGAGAAACATGAAACAAACTGATTTTCTAATAATGAAGAGAGAG 24376
296 CCGATTCATGCTACCGCTTATGATTTCCGCGAGCACTTAACCTGGAACAGCGGA 355
24375 CCGATTCATGCTACCGCTTATGATTTCCGCGAGCACTTAACCTGGAACAGCGG 24316
356 GTTGAATGATTTAGTCCGATTCATGGAATGCTGCTCGGCTTGATTTCACT 415
24315 GTTGAATGATTTAGTCCGATTCATGGAATGCTGCTCGGCTTGATTTCACT 24256
416 GTTGTGTGACAGTTGCGGACATGATCCATCAAAACCGTTTAAAGGGGTGCGCG 475
24255 GTTGTGTGACAGTTGCGGACATGATCCATCAAAACCGTTTAAAGGGGTGCGCG 24196
476 AATACCTTATGTGACAG 494
24195 AATACCTTATGTGACAG 24177

RESULT 13
AX572872/c AX572872 6886 bp DNA 1linear PAT 29-NOV-2002
LOCUS
DEFINITION Sequence 20 from Patent WO02057476.

ACCESSION AX572872
VERSION AX572872.1 GI:26004958
KEYWORDS
SOURCE synthetic construct
ORGANISM synthetic construct
REFERENCE
1
AUTHORS Hermann, T., Patterson, T.A., Pero, J.G., Yocum, R.R., Balderius, K.U.
and Beck, C.
TITLE Methods and microorganisms for the production of
3-(2-hydroxy-3-methyl-butyrylamino)-propionic acid (hmbpa)
JOURNAL Patent: WO 02057476-A 20 25-JUL-2002;
FEATURES
location/Qualifiers
SOURCE
1..6886
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="vector-PAN624"

ORIGIN

Query Match 43.8%; Score 216.2; DB 6; Length 6886;

Best Local Similarity 98.6%; Pred. No. 1e-47; Mismatches 3; Indels 0; Gaps 0;

Matches 218; Conservative 0;

1 GCTAATGTTGGTACAGCCCGTGAATTTGTATATCTTCATTGGGCGAGTATCGCTG 60
2909 GCTAATGTTGGTACAGCCCGTGAATTTGTATATCTTCATTGGGCGAGTATCGCTG 2850
61 CGAAGTGCATCTATTTAATAATAGATAGACATTGACAGCTGCTGCTTATCCAAAAA 120
2849 CGAAGTGCATCTATTTAATAATAGATAGACATTGACAGCTGCTGCTTATCCAAAAA 2790
121 GGAAGTGCAGAGAGGATGAACTCGCCGAACTTTAGAAAGTGAAGATCTTCTGTTG 180
2789 GGAAGTGCAGAGAGGATGAACTCGCCGAACTTTAGAAAGTGAAGATCTTCTGTTG 2730
181 TACGGAAGGTTTGTGCTTGCAAGAAACGCGAGATC 221
2729 TACGGAAGGTTTGTGCTTGCAAGAAACGCGAGATC 2689

RESULT 14

AX573224/c AX573224 6886 bp DNA 1linear PAT 29-NOV-2002

LOCUS
DEFINITION Sequence 20 from Patent WO02057474.

AX573224

VERSION
AX573224.1 GI:26005137

KEYWORDS

SOURCE

ORGANISM

synthetic construct
synthetic construct
other sequences; artificial sequences.

REFERENCE

AUTHORS

1 Hermann, T., Patterson, T.A., Pero, J.G., Yocum, R.R., Balderius, K.U.
and Beck, C.

TITLE Processes for enhanced production of pantothenate

JOURNAL Patent: WO 02057474-A 20 25-JUL-2002;

OMNIGENE BIOPRODUCTS, INC. (US)

FEATURES

SOURCE

location/Qualifiers

1..6886

/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="vector-PAN624"

ORIGIN

Query Match 43.8%; Score 216.2; DB 6; Length 6886;

Best Local Similarity 98.6%; Pred. No. 1e-47; Mismatches 3; Indels 0; Gaps 0;

Matches 218; Conservative 0;

1 GCTAATGTTGGTACAGCCCGTGAATTTGTATATCTTCATTGGGCGAGTATCGCTG 60
2909 GCTAATGTTGGTACAGCCCGTGAATTTGTATATCTTCATTGGGCGAGTATCGCTG 2850

Oy	61	GGAACTGCAGCCATATATTAATTAATAGATAGACACTTGCACAGCTGCCCTTGAATCCAAAAA	120
Db	2849	CGAACTGCAGCCATATATTAATTAATAGATAGACACTTGCACAGCTGCCCTTGAATCCAAAAA	2790
Oy	121	GGACTGGACAGAGGAGTAAACTGCGCAACTTAGAGAAAGTGAAGATCCTTCCTGTG	180
Db	2789	GGACTGGACAGAGGAGTAAACTGCGCAACTTAGAGAAAGTGAAGATCCTTCCTGTG	2730
Oy	181	TACCGAAGCTTTTGGCTTGCAGAGAAAAACGCGACATC	221
Db	2729	TACCGAAGCTTTTGGCTTGCAGAGAAAAACGCGAATTC	2689
RESULT 15			
BACJ00C			
LOCUS		6540 bp	DNA
DEFINITION		1	linear
			BCT 26-MAR-1996
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
" JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
FEATURES			
source			
gene			
CDS			
terminator			
gene			
CDS			
gene			
CDS			

gene
CDS
3229..4302
/gene="daps"
/function="lysine biosynthesis"
/note="28.2% of identical aminoacids with the Escherichia coli dihydrodipicolinate reductase; putative"
/codon_start=1
/transl_table=1
/product="dihydrodipicolinate reductase"
/protein_id="AAA92874.1"
/db_xref="GI:755603"
/db_xref="GI:755603"
/translation="MSNSTIKLVLAGRGNGEAVKLARTPFEDLVGALIDHTYDOQ
KLSDMVPESDAFIYTDIHACFTPTPDVLAIDTPPIGSGVHTKIALHSEVPRVGT
GFSEDALELTLEKTEKIGALIIAPNELKIGLIMKRSKMAVNFEDVEVLIELHDOCK
LDAESGTALKTAEIMISEIKRKKQGGDEKEILPGARGAONGIRLRLRIHDSVNR
ASPMGVALSTEQMKIDQLVYGLNIIID"
2117..2530
/gene="jofF"
/translation="MKIALIADHKKKQDMVOFTTAYRDLKNHDLVATGTGAKIHBA
TSLQIERQSSPLGGDOQIGALIANALDIYILRDLPLTAQPHPEPVSALIRCDVYS
IPLATNGTAELIVRTLDGVEFPFDLIRGSEPNV"
2523..3197
/gene="jofG"
/translation="MYNADVLAFGAHSDDVIEIGNGTIAFVQSEKVMICDLTEAL
SSNGVSLRKEEAALARIIGADRIQULTLPDRLLISDQAINISIVVILCRPAVF
MPEKGRDRPDRGNAALVEEALFSAIGIHKDEKSLPAHKVSKVYMYINGFQPDV
IDISLTLEAKKQSLNAYKQPIPSKDSVSTPLTNGYIEIVAREKLYGKGAQVEYAEV
SEPNQC"
3229..4302
/gene="jofH"
/translation="WRKLGKOLAKGHEIHFTSSIPRLNTYVPHNHPHEVNOVA
VFKEPYDITLASKIAEVAERENLDITIAHIALPHACAYLAKOMLKRNIQVYTLTG
TDIVLGYDPLKDLIRPALESSDRAVASSALAAETVYLIRPEKKEITLYNFDDEV
YKTKATAIKKKHGILPDEKVIHVSNFRKVRQVDIRVPRNAGTKTALLVGGGGL
PEKSNAGCALIKRYGLEDVLMGRQDVEDLYSISDLKLSEKSGVLVLEMAAGG
VPCIGTNGGIPVYIKNNVSGFLVDGVDVTAATARAASILEDQLSNRFKAIMELE
NEFSKATVSGYEOTIADLAEPE"
4307..5440

CDS

```

/gene="paps"
4307..5440
/gene="paps"
/note="38.9% of identical aminoacids with the Escherichia
coli pcbB gene product, encoding poly(A) polymerase;
putative"
/codon_start=1
/transl_table=11
/product="polyA polymerase"
/protein_id="AA92878.1"
/db_xref="gi:755607"
/translation="MEKVFIALPVLRLIEAGHOAYFVGAVRDSYMKRTIGPDVIA
TDAPDOVERLFOITVDYKDESDYDPRPSEYQFISLEEDLRDLITINAMATA
DGKLDYFGGKDIIDKVIPTVGKEDRFRSDALRMLRAYFMSOLGPTLSPEIEAI
AKESLSLHVSVERKTIIEFKLQGRASRQALQTLIOTRLYEELPGYHRENLSIS
EPPFSLTSREELMAALINIGIVLKDAPFLKAKLPGKYKEAIIHADTFQSIDA
MTMYRAGKKALISAQISQLRNENKDEKDKDIQYAYQNLPIKSLKDLDTGKDLA
LRRPAGKWSVEELQWIEQAVVTGKLSNQKHIEEWLTCGQH"
5425..6342
/gene="bira"
5425..6342
/function="bifunctional biotin operon repressor"
/note="26.7% identity to the Escherichia coli bifunctional
biotin operon repressor"
/codon_start=1
/evidence=experimental
/transl_table=11
/product="biotin [acetyl-CoA carboxylase] ligase"
/protein_id="AA92879.1"
/db_xref="gi:755608"
/translation="MRSTLRKDLIELRSQAGNEFISGKISDALGCSRVAWKIEEL
RKEGYEAVRKRGLIKPKGLSESIIRGLTEVWGQHLIYHVLSTQKTAHEL
ANNNAPBEGTLENGVMSLILRPDIPLQKTPQLTLAAVAVQGIIEAAGIQTDIKP
NDILNGKTVGILTEMQAEDRVSIVIGINWQOPNDPDELKDIATLSQAAQ
EKIDRAGVIOHILCFEKRVRDVMTHGFTPIKLWESYALGIGTMMRARTLNGTFYQK
ALGIDDEGVLLTNEGIKIYSADIEIG"
6501..6533
/note="putative"

```

terminator

```

/note="putative"

```

ORIGIN

```

Query Match      41.1%; Score 203; DB 1; Length 6540;
Best Local Similarity 100.0%; Pred. No. 3.9e-44;
Matches 203; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCTAATGTTGGTACAAAGCCGTTGATTTGGTATTACTCCATTGGCAGTATGCCTG 60
DB 6338 GCTAATGTTGGTACAAAGCCGTTGATTTGGTATTACTCCATTGGCAGTATGCCTG 6397
QY 61 CGAAGTGCACCTATTATTAAATAGATAGACATTGCAGAGTCTGCCTTGATCCAAAAA 120
DB 6398 CGAAGTGCACCTATTATTAAATAGATAGACATTGCAGAGTCTGCCTTGATCCAAAAA 6457
QY 121 GGAAGTGCAGAGGATGAAGCTGCCGAATTTAGAAAGTGAAGTCCCTTCGTTG 180
DB 6458 GGAAGTGCAGAGGATGAAGCTGCCGAATTTAGAAAGTGAAGTCCCTTCGTTG 6517
QY 181 TAAAGGAAGTTTGGCTTGC 203
DB 6518 TAAAGGAAGTTTGGCTTGC 6540

```

Search completed: September 8, 2005, 21:13:00
 Job time : 2370 secs

This Page Blank (uspic)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using SW model

Run on: September 8, 2005, 16:46:49 / Search time 7320 Seconds
(without alignments)
399.502 Million cell updates/sec

Title: US-09-925-824A-1

Perfect score: 494
Sequence: 1 gctaatgtgtgtgacacg.....gaataccttattgtgacag 494

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database:

N_Geneseq_16Dec04:
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001s:*
5: geneseqn2002s:*
6: geneseqn2003as:*
7: geneseqn2003bs:*
8: geneseqn2003cs:*
9: geneseqn2003ds:*
10: geneseqn2004as:*
11: geneseqn2004bs:*
12: geneseqn2004cs:*
13: geneseqn2004ds:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	494	100.0	494	6	ABA95293 B. subtilis
2	485	98.2	2363	4	AAS02314 B. subtilis
3	258	52.2	7381	4	AAS02335 Plasmid P
4	257	52.0	6725	6	AAL48885 B. subtilis
5	257	52.0	6725	6	AAL41229 Pantocthen
6	257	52.0	8503	4	AAS02334 Plasmid P
7	257	52.0	10801	4	AAS02333 Plasmid P
8	244	49.4	831	4	AAS00988 B. subtilis
9	224.6	45.5	10181	12	ADN36367 Plasmid P
10	216.2	43.8	6886	6	AAL48883 B. subtilis
11	216.2	43.8	6886	6	AAL41227 Pantocthen
12	151.2	30.6	834	6	ABK73715 Bacillus
13	139	28.1	837	6	ACA21619 Prokaryot
14	137	27.7	1613	6	ABQ70448 Listeria
15	137	27.7	110000	6	ABQ69245-20 Continuation (21 o
16	137	27.7	319630	6	ABQ67194 Listeria
17	133.8	27.1	110000	6	ABA03041-19 Continuation (20 o
18	128.8	26.1	834	8	ACA36655 Prokaryot
19	117.2	23.7	1880	4	AHH5062 S. epider
20	111.6	22.6	816	8	ACA46771 Prokaryot

21	111.6	22.6	825	6	ABN92106	ABN92106 Staphyloc
22	111.6	22.6	825	13	ADSO1749	Adso1749 Staphyloc
23	105.8	21.4	865	8	ACA22421	ACA22421 Prokaryot
24	103.2	20.9	792	4	AAH53012	AAH53012 S. epider
25	99.6	20.2	4549	2	AAV74485	AAV74485 Staphyloc
26	94	19.0	816	8	AAS52191	AAS52191 Staphyloc
27	94	19.0	816	8	ACF74215	ACF74215 Staphyloc
28	92.4	18.7	825	8	ACA20255	ACA20255 Prokaryot
29	87.4	17.7	825	8	ACA32887	ACA32887 Prokaryot
30	87.4	17.7	825	8	ACA33554	ACA33554 Prokaryot
31	87.4	17.7	825	4	AAS53108	AAS53108 Enterococ
32	87.4	17.7	828	10	ADCC90732	ADCC90732 E. faeciu
33	87.4	17.7	828	10	ADH85238	ADH85238 Enterococ
34	85.6	17.3	792	4	AAS54927	AAS54927 Staphyloc
35	85.4	17.3	491	4	AAS50306	AAS50306 Staphyloc
36	85.4	17.3	491	8	ACA17539	ACA17539 Prokaryot
37	85.4	17.3	816	8	ACA47659	ACA47659 Prokaryot
38	78	15.8	3301	4	AAH54845	AAH54845 S. epider
39	74.4	15.1	831	8	ACA27876	ACA27876 Prokaryot
40	74.2	15.0	855	9	ADA32533	ADA32533 DNA encod
41	72.8	14.7	807	8	ACA21162	ACA21162 Prokaryot
42	72	14.6	349980	5	AAH41224	AAH41224 Pyrococcu
43	71.8	14.5	825	8	ACA28486	ACA28486 Prokaryot
44	71.6	14.5	825	8	ACA28877	ACA28877 Prokaryot
45	70	14.2	789	8	ACA41207	ACA41207 Prokaryot

ALIGNMENTS

RESULT 1
ABA95293 standard; DNA, 494 BP.

ABA95293; 10-JUN-2002 (first entry)

B. subtilis panB promoter sequence.

Bacterium; tetrahydrofolate; THF; panB; promoter; para-aminobenzoic acid; PABA; antibacterial; ds.

Bacillus subtilis.

WO200214559-A2.

21-FEB-2002.

10-AUG-2001; 2001WO-US041665.

11-AUG-2000; 2000US-0224925P.

(MILL-) MILLENNIUM PHARM INC.

Murphy C; WPI; 2002-269209/31.

Identifying inhibitor of bacterial tetrahydrofolate biosynthesis for treating bacterial infection, by contacting a cell having ketopantoate hydroxymethyltransferase promoter with an agent and measuring promoter activity.

Claim 3; Page 14-15; 31pp; English.

The invention relates to a method of determining whether a test compound is an inhibitor of bacterial tetrahydrofolate (THF) biosynthesis. The method involves contacting a bacterial cell with the test compound, where the cell contains a promoter, the activity of which is increased in the presence of a compound that inhibits THF biosynthesis, and measuring activity of the promoter. The method is useful for determining whether a test compound is an inhibitor of bacterial THF biosynthesis, and also for

determining whether a test compound is an antibacterial agent. The compounds that inhibit THF biosynthesis can be identified by their ability to increase the activity of the B. subtilis panB promoter. The inhibition of THF biosynthesis is detected as inhibition of para-aminobenzoic acid (PABA) uptake into cells and the inhibition is measured in a biochemical assay with a cell extract for an enzyme activity which is required for THF biosynthesis. The inhibitor identified by the method is useful as a lead compound for preparing antibacterial agents.

Compositions comprising the antibacterial agent is useful for treating a bacterial infection caused by S. pneumoniae, S. pyogenes, S. agalactiae, S. endocarditis, S. faecium, S. sanguis, S. viridans, S. haemolyticus in a mammal, in particular a human. The compounds can also be used to treat infection of gram negative bacteria e.g., Shigella, E. coli, Klebsiella and Yersinia. The present sequence represents the B. subtilis panB promoter sequence

Sequence 494 BP; 147 A; 95 C; 122 G; 130 T; 0 U; 0 Other;

Query Match 100.0%; Score 494; DB 6; Length 494;
Best Local Similarity 100.0%; Pred. No. 3.6e-143;
Matches 494; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GCTATGTTGGTGTCAAGCCCGTTGATTTGGTATCTTCCATTGGGCGATGCTCG 60
DB 1 GCTATGTTGGTGTCAAGCCCGTTGATTTGGTATCTTCCATTGGGCGATGCTCG 60
OY 61 CGAATCGACCTATTAATAATAGATGACATTCGACGAGCTGCTGATCCAAAAA 120
DB 61 CGAATCGACCTATTAATAATAGATGACATTCGACGAGCTGCTGATCCAAAAA 120
OY 121 GGACTGGGACAGAGGATGAAAATCGCGAATTTAGAAAGTGAAGAAATCTTCGTTG 180
DB 121 GGACTGGGACAGAGGATGAAAATCGCGAATTTAGAAAGTGAAGAAATCTTCGTTG 180
OY 181 TAAACGAAAGCTTTTGGCTTCGACAGAAAAACGACAGTCACTCTCTTAAACATGAGG 240
DB 181 TAAACGAAAGCTTTTGGCTTCGACAGAAAAACGACAGTCACTCTCTTAAACATGAGG 240
OY 241 AGGAGAAAACATGAAAACAAAATCTGATTTCTAATAATGAAGAGTCTGAAGAACGAT 300
DB 241 AGGAGAAAACATGAAAACAAAATCTGATTTCTAATAATGAAGAGTCTGAAGAACGAT 300
OY 301 TGTCAATGTCGACCGCTTATGATTTATCCGACCTAAACTTGTGAACAAGCGAGTTGA 360
DB 301 TGTCAATGTCGACCGCTTATGATTTATCCGACCTAAACTTGTGAACAAGCGAGTTGA 360
OY 361 CATGATTTTAAATGCGGTGATTCCTTGAATGCTGCTCGGCTTGAATCACTGTGG 420
DB 361 CATGATTTTAAATGCGGTGATTCCTTGAATGCTGCTCGGCTTGAATCACTGTGG 420
OY 421 TGTGACAGTGGGACATGATCATATCAAAAACCGTTAAAGGGGTGCGCGAATAC 480
DB 421 TGTGACAGTGGGACATGATCATATCAAAAACCGTTAAAGGGGTGCGCGAATAC 480
OY 481 CTTTATTTGTGACAG 494
DB 481 CTTTATTTGTGACAG 494

RESULT 2
AAS02314
ID AAS02314 standard; DNA; 2363 BP.
XX AAS02314;
AC
XX 18-JUL-2001 (first entry)
DT
XX
DE B. subtilis panBCD operon.
XX
KW PanBCD operon; PanC; PanB; pantothenate biosynthesis; vitamin B5;
KM nutritional supplement; panto-compound; pantoate; ds.
XX
OS Bacillus subtilis.

XX Key Location/Qualifiers
FH CDS 242..1075
FT CDS /*cag= a
FT CDS /product= "PanB"
FT CDS /cag= b
FT CDS /product= "Panc"
FT CDS 1939..2332
FT CDS /*cag= c
FT CDS /product= "Pand"

WO200121772-A2.
29-MAR-2001.
21-SEP-2000; 2000MO-US025993.
21-SEP-1999; 99US-00400494.
07-JUN-2000; 2000US-0210072P.
28-JUL-2000; 2000US-0221836P.
24-AUG-2000; 2000US-0227860P.
(OMNI-) OMNIGENE BIOPRODUCTS.
Yocum RR, Paterson TA, Hermann T, Pero JG;
WPI: 2001-218644/22.
P-PSDB; AAU01244, AAU01245, AAU01246.
New recombinant microorganism which overexpresses a Bacillus subtilis pantothenate biosynthetic enzyme, useful for the high yield production of panto-compounds such as pantothenate and pantoate.
Claim 86; Page 198-201; 292pp; English.
The sequence represents the B. subtilis PanBCD operon encoding PanB, PanC and PanD, enzymes of the pantothenate biosynthetic pathway. Pantothenate, also known as vitamin B5, is used as a nutritional supplement in mammals and humans. The invention concerns methods of producing recombinant microorganisms overexpressing at least one Bacillus subtilis pantothenate biosynthetic enzyme. The microorganisms and methods of producing them are useful for producing a panto-compound such as pantothenate or pantoate, which is a nutritional requirement for livestock and humans. The methods are also useful for the identification of pantothenate kinase modulators. Panto-compounds are produced at a significantly higher yield than prior art methods and can be produced independent of the need to feed precursors which decreases expense

Sequence 2363 BP; 730 A; 451 C; 593 G; 589 T; 0 U; 0 Other;

Query Match 98.2%; Score 485; DB 4; Length 2363;
Best Local Similarity 100.0%; Pred. No. 4.6e-140;
Matches 485; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 10 TTGGTACAGCCCGTTGATTTGGTATCTTCCATTGGGCGATGCTCGCAACTGCA 69
DB 1 TTGGTACAGCCCGTTGATTTGGTATCTTCCATTGGGCGATGCTCGCAACTGCA 60
OY 70 CCTATTATAATATGATGATGACATTCGACAGTCTGCTTGAATCAAAAAGACTGGGA 129
DB 61 CCTATTATAATATGATGATGACATTCGACAGTCTGCTTGAATCAAAAAGACTGGGA 120
OY 130 CAGAGGATGAAATCGCCGCACTTTGAAAGTGAAGTGAAGTCTTCTGTTAAGGAG 189
DB 121 CAGAGGATGAAATCGCCGCACTTTGAAAGTGAAGTGAAGTCTTCTGTTAAGGAG 180
OY 190 GTTTTGGCTTGGCAGAGAAAACGACAGATCATCTCTCTAACAATGAGAGAGAAA 249
DB 181 GTTTTGGCTTGGCAGAGAAAACGACAGATCATCTCTCTAACAATGAGAGAGAAA 240
OY 250 CATGAAAACAAACTGGAATTTTCTAATAATGAAGAGTCTGAAGAACGATTTGTATGCT 309

	SQ	Sequence 7381 BP; 1899 A; 1823 C; 1601 G; 2058 T; 0 U; 0 Other;
		Query Match 52.2%; Score 258; DB 4; Length 7381;
		Best Local Similarity 100.0%; Pred. No. 3.6e-69;
		Matches 258; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
CC	OY	237 GAGGAGGAGAAAACATGAACAACAAACTCGATTCTTCTTAATAAATGAAGAGTCGTAAAGAC 296
CC	DB	3652 GAGGAGGAGAAAAACATGA AAAACA AAACCGATTCTTCTAAAAATGAAGAGTCGTAAAGAC 35939
CC	OY	297 CGATTGTCATCCTGACCGCCTTAGTATATATTCGGGACGCTAAACCTTGCTGAACAAGCGGAG 356
CC	DB	3592 CGATTGTCATCCTGACCGCCTTAGTATATTCGGGACGCTAAACCTTGCTGAACAAGCGGAG 35333
CC	OY	357 TTGACATGATTTTATAGTCGGTATCATCTTGGAATGTCGTCCTCGGCCTTGATTCAACTG 416
CC	DB	3552 TTGACATGATTTTATAGTCGGTATCATCTTGGAATGTCGTCCTCGGCCTTGATTCAACTG 34737
CC	OY	417 TCGGTGTGACAGTTGCGGACATGATTCATATCAAAAGCCGTTAAAGGGGTGCGCCGA 476
CC	DB	3412 TCGGTGTGACAGTTGCGGACATGATTCATATCAAAAGCCGTTAAAGGGGTGCGCCGA 34131
CC	OY	477 ATACTTTATTTGTGACAG 494
CC	DB	3412 ATACTTTATTTGTGACAG 3395
CC	RESULT 4	
CC	AAL48885	
CC	ID	AAL48885 standard; DNA; 6725 BP.
CC	XX	
CC	AC	AAL48885;
CC	XX	
CC	DT	24-OCT-2002 (first entry)
CC	DE	B subtilis PanB gene vector PAN636.
CC	XX	
CC	KW	3-(2-hydroxy-3-methyl-butylamino)-propionic acid; HMBPA; alpha-HIV;
CC	KM	alpha-hydroxylisovalerate; panicochinate biosynthesis;
CC	KW	cardiovascular disease; HMG CoA reductase; hypercholesterolaemia;
CC	KW	atherosclerosis; skin aging; PanB; Pand; PanE1; PanE2; vector; ds.
CC	XX	
CC	OS	Bacillus subtilis.
CC	OS	Synthetic.
CC	XX	
CC	PN	WO200257476-A2.
CC	XX	
CC	PD	25-JUL-2002.
CC	XX	
CC	PF	19-JAN-2002; 2002WO-US001887.
CC	PR	19-JAN-2001; 2001US-0263053P.
CC	XX	
CC	PA	(OMNI-) OMNIGENE BIOPRODUCTS INC.
CC	PI	
CC	XX	Hermann T, Patterson TA, Pero JG, Yocum RR, Balidenius K, Beck C;
CC	DR	WFI; 2002-608383/65.
CC	PT	Preparation of 3-(2-hydroxy-3-methyl-butylamino)-propionic acid useful
CC	PT	for synthesizing hydroxy methyl glutarate COA reductase inhibitors,
CC	PT	involves culturing a microorganism under suitable conditions and
CC	XX	detecting or isolating product.
CC	XX	
CC	PS	Example 4; Page 72-74; 80pp; English.
CC	XX	The present invention relates to a method of producing 3-(2-hydroxy-3-
CC	CC	methyl-butylamino)-propionic acid (HMBPA), which involves culturing a
CC	CC	microorganism under conditions such that HMBPA is produced, and detecting

or isolating the HMBPA produced by the microorganism. These organisms may have modified pantothenate biosynthesis enzymes. The method can be used to produce HMBPA, which can then be used to synthesize inhibitors of hydroxy methyl glutarate (HMG) CoA reductase, which are useful for treating hypercholesterolemia, coronary atherosclerosis progression, and to reduce risk of cardiovascular event in patients at risk. Alpha-HIV is also produced, which is useful in the prevention of aging of skin, and to treat skin disorders such as age spots, skin lines, wrinkles, photoaging and aging. The present sequence is a vector used in the exemplification of the invention

Sequence 6725 BP; 1773 A; 1541 C; 1668 G; 1743 T; 0 U; 0 Other;

Query Match 52.0%; Score 257; DB 6; Length 6725;

Best Local Similarity 100.0%; Pred. No. 7e-69; Mismatches 0; Indels 0; Gaps 0;

Matches 257; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 238 AGAGAGAGAAAAACATGAAAAACAAACCTGATTTCTTAAAAATGAGAGCTGAGAAC 297
    |||
DB 325 AGAGAGAGAAAAACATGAAAAACAAACCTGATTTCTTAAAAATGAGAGCTGAGAAC 384
    |||
QY 298 GATTGTCATGCTGACCGCTTATGATTATCCGGACGCTAACTTGCTGAACAAAGCGGAGT 357
    |||
DB 385 GATTGTCATGCTGACCGCTTATGATTATCCGGACGCTAACTTGCTGAACAAAGCGGAGT 444
    |||
QY 358 TGACATGATTTTAACTCGGTGATTCACCTTGAATGCTGCTCGGCTTGAATCAACTGT 417
    |||
DB 445 TGACATGATTTTAACTCGGTGATTCACCTTGAATGCTGCTCGGCTTGAATCAACTGT 504
    |||
QY 418 CGGTGTGACAGTTGCGGACATGATCCATCATACAAAGCCGTTAAAGGGGTGCGCGAA 477
    |||
DB 505 CGGTGTGACAGTTGCGGACATGATCCATCATACAAAGCCGTTAAAGGGGTGCGCGAA 564
    |||
QY 478 TACCTTTATTTGTGACAG 494
    |||
DB 565 TACCTTTATTTGTGACAG 581
    |||
```

RESULT 5

AAAL41229

ID AAL41229 standard; DNA; 6725 BP.

AC AAL41229;

DT 30-OCT-2002 (first entry)

DE Pantothenate related vector PAN636 SEQ ID No 22.

XX 3-(2-hydroxy-3-methyl-butylamino)-propionic acid; HMBPA; PanB; PanB1;
XX pantothenate composition; deregulated pantothenate biosynthetic pathway;
XX deregulated isoleucine-valine; ilv biosynthetic pathway; PanB2; IlvC; ds.
XX unidentified.

OS WO200257474-A2.

PD 25-JUL-2002.

PP 19-JAN-2002; 2002WO-US001842.

PR 19-JAN-2001; 2001US-0262995P.

PA (OMNI-) OMNIGENE BIOPRODUCTS INC.

PI Hermann T, Patterson TA, Pero JG, Yocum RR, Baldenius K, Beck C;

DR WPI; 2002-636526/68.

PT Preparation of 3-(2-hydroxy-3-methyl-butylamino)-propionic acid-free
PT pantothenate composition by culturing microorganism having deregulated
PT pantothenate biosynthetic pathway and modified pantothenate biosynthetic
PT enzyme activities.

PS Example 4; Page 80-82; 93pp; English.

XX The invention relates to a method for the preparation of 3-(2-hydroxy-3-
XX methyl-butylamino)-propionic acid (HMBPA)-free pantothenate composition
CC by culturing microorganism having deregulated pantothenate biosynthetic
CC pathway and deregulated isoleucine-valine (ilv) biosynthetic pathway. The
CC microorganism has regulated activities of PanB; PanB1; PanB2; IlvC; PanB
CC and PanB1; PanB and PanB2; or, PanB and IlvC. The methods reduce HMBPA
CC (byproduct) production and provide increasing yields and purity of the
CC desired product. This polynucleotide sequence represents a vector
CC sequence relating to the pantothenate compositions of the invention

Sequence 6725 BP; 1773 A; 1541 C; 1668 G; 1743 T; 0 U; 0 Other;

Query Match 52.0%; Score 257; DB 6; Length 6725;

Best Local Similarity 100.0%; Pred. No. 7e-69; Mismatches 0; Indels 0; Gaps 0;

Matches 257; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 238 AGAGAGAGAAAAACATGAAAAACAAACCTGATTTCTTAAAAATGAGAGCTGAGAAC 297
    |||
DB 325 AGAGAGAGAAAAACATGAAAAACAAACCTGATTTCTTAAAAATGAGAGCTGAGAAC 384
    |||
QY 298 GATTGTCATGCTGACCGCTTATGATTATCCGGACGCTAACTTGCTGAACAAAGCGGAGT 357
    |||
DB 385 GATTGTCATGCTGACCGCTTATGATTATCCGGACGCTAACTTGCTGAACAAAGCGGAGT 444
    |||
QY 358 TGACATGATTTTAACTCGGTGATTCACCTTGAATGCTGCTCGGCTTGAATCAACTGT 417
    |||
DB 445 TGACATGATTTTAACTCGGTGATTCACCTTGAATGCTGCTCGGCTTGAATCAACTGT 504
    |||
QY 418 CGGTGTGACAGTTGCGGACATGATCCATCATACAAAGCCGTTAAAGGGGTGCGCGAA 477
    |||
DB 505 CGGTGTGACAGTTGCGGACATGATCCATCATACAAAGCCGTTAAAGGGGTGCGCGAA 564
    |||
QY 478 TACCTTTATTTGTGACAG 494
    |||
DB 565 TACCTTTATTTGTGACAG 581
    |||
```

RESULT 6

AAAS02334/C

ID AAS02334 standard; DNA; 8503 BP.

AC AAS02334;

DT 06-AUG-2003 (revised)

DT 18-JUL-2001 (first entry)

DE Plasmid PAN004 carrying the B. subtilis PanBCD operon.
XX PanBCD operon; pantothenate biosynthesis; cyclic; circular; PAN004;
XX vitamin B5; nutritional supplement; panto-compound; pantoate; P26; RBS1;
XX ribosome binding site; ds.

OS Bacteriophage.

OS Escherichia coli.

OS Bacillus subtilis.

OS Chimeric.

PP WO200121772-A2.

PD 29-MAR-2001.

PP 21-SEP-2000; 2000WO-US025993.

PR 21-SEP-1999; 99US-00400494.

PR 07-JUN-2000; 2000US-0210072P.

PR 28-JUL-2000; 2000US-0221836P.

PR 24-AUG-2000; 2000US-0227860P.

PA (OMNI-) OMNIGENE BIOPRODUCTS.

PI Yocum RR, Patterson TA, Hermann T, Pero JG;


```
XX OS Bacillus subtilis.
XX FH Key Location/Qualifiers
XX FT CDS 1..831
XX FT /*tag= a
XX FT /product= "Ketopantoate hydroxymethyltransferase"
XX FT /partial
XX FT /note= "No stop codon"
XX FN WO200121772-A2.
XX PD 29-MAR-2001.
XX PF 21-SEP-2000; 2000WO-US025993.
XX PR 21-SEP-1999; 99US-00400494.
XX PR 07-JUN-2000; 2000US-0210072P.
XX PR 28-JUL-2000; 2000US-0221836P.
XX PR 24-AUG-2000; 2000US-0227860P.
XX PA (OMNI-) OMNIGENE BIOPRODUCTS.
XX PI Yocum RR, Patterson TA, Hermann T, Pero JG;
XX DR WPI; 2001-218644/22.
XX DR P-PSDB; AAU01244.
XX PS New recombinant microorganism which overexpress a Bacillus subtilis
XX PT pantothenate biosynthetic enzyme, useful for the high yield production of
XX PT panto-compounds such as pantothenate and pantoate.
XX PS Claim 86; Page 169-170; 292pp; English.
XX CC The sequence, panB, encodes B. subtilis ketopantoate
XX CC hydroxymethyltransferase, an enzyme of the pantothenate biosynthetic
XX CC pathway. Pantothenate, also known as vitamin B5, is used as a nutritional
XX CC supplement in mammals and humans. The invention concerns methods of
XX CC producing recombinant microorganisms overexpressing at least one Bacillus
XX CC subtilis pantothenate biosynthetic enzyme. The microorganisms and methods
XX CC of producing them are useful for producing a panto-compound such as
XX CC pantothenate or pantoate, which is a nutritional requirement for
XX CC livestock and humans. The methods are also useful for the identification
XX CC of pantothenate kinase modulators. Panto-compounds are produced at a
XX CC significantly higher yield than prior art methods and can be produced
XX CC independent of the need to feed precursors which decreases expense
XX CC
XX SQ Sequence 831 BP; 249 A; 166 C; 219 G; 197 T; 0 U; 0 Other;
XX
XX Query Match 49.4%; Score 244; DB 4; Length 831;
XX Best Local Similarity 100.0%; Pred. No. 3,1e-65;
XX Matches 244; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 251 ATGAAACAAACCTGATTTTCTAAATAAGAGAGTGAAGAACCGATTGTCATGCTG 310
XX DB 1 ATGAAACAAACCTGATTTTCTAAATAAGAGAGTGAAGAACCGATTGTCATGCTG 60
XX QY 311 ACCGTTATGATTAATCCGGACCTAAACTGTGTGAACAAGCGGAGTTGACATGATTTTA 370
XX DB 61 ACCGTTATGATTAATCCGGACCTAAACTGTGTGAACAAGCGGAGTTGACATGATTTTA 120
XX QY 371 GTCCGTGATTCACCTGGAATGTCGTCTCCGCGCTTGAATCAACTGTGCGTGTGACATT 430
XX DB 121 GTCCGTGATTCACCTGGAATGTCGTCTCCGCGCTTGAATCAACTGTGCGTGTGACATT 180
XX QY 431 GCGGACATGATCCATCATACAAACCGTTAAAGGGGTGCGCCGAATACCTTTATTGTG 490
XX DB 181 GCGGACATGATCCATCATACAAACCGTTAAAGGGGTGCGCCGAATACCTTTATTGTG 240
XX QY 491 ACAAG 494
XX DB 241 ACAAG 244
```

```
XX RESULT 9
XX ID ADN36367 standard; DNA; 10181 BP.
XX AC ADN36367;
XX DT 15-JUL-2004 (first entry)
XX DE Plasmid pDX17R for expression of the Bacillus subtilis yaadE sequence.
XX KW B6 vitamin; Yaad; Yaaf; Bpd; PdxA; PdxJ; PdxF; PdxB; PdxH; Dxs;
XX KW pyridoxine; pyridoxal; pyridoxamine; ds.
XX OS Bacillus subtilis.
XX FN WO2004035010-A2.
XX PD 29-APR-2004.
XX PF 21-MAR-2003; 2003WO-US008880.
XX PR 22-MAR-2002; 2002US-0367089P.
XX PR 25-MAR-2002; 2002US-0367863P.
XX PR 29-MAR-2002; 2002US-0368618P.
XX PR 03-MAR-2003; 2003US-00366618.
XX PA (BADI ) BASF AG.
XX PI Yocum RR, Williams MK, Pero JG;
XX DR WPI; 2004-365085/34.
XX PS Producing B6 vitamin such as pyridoxine, pyridoxal or pyridoxamine,
XX PT involves culturing organism with increased yaad and/or yaaf activity as
XX PT compared to parent organism.
XX PS Example 4; SEQ ID NO 8; 75pp; English.
XX CC The invention relates to a method of producing (M1) a B6 vitamin, which
XX CC involves culturing an organism (I) with increased yaad and/or yaaf
XX CC activity, or Bpd, PdxA, PdxJ, PdxF, PdxB, PdxH and/or Dxs activity as
XX CC compared to the parent organism, or a microorganism that overexpresses
XX CC Bacillus B6 vitamin biosynthetic gene, under conditions such that B6
XX CC vitamin is produced. (M1) is useful for producing B6 vitamin such as
XX CC pyridoxine, pyridoxal and pyridoxamine. This sequence represents the
XX CC plasmid pDX17R which contains the Bacillus subtilis yaadE operon used to
XX CC express these proteins.
XX SQ Sequence 10181 BP; 2655 A; 2363 C; 2041 G; 3122 T; 0 U; 0 Other;
XX
XX Query Match 45.5%; Score 224.6; DB 12; Length 10181;
XX Best Local Similarity 92.5%; Pred. No. 1e-58;
XX Matches 236; Conservative 0; Mismatches 19; Indels 0; Gaps 0;
XX
XX QY 1 GCTAATGTGTGATCAAGCCCGTGAATTTGGTATTAATTCATTTGGGAGATGCGCTG 60
XX DB 3570 GCTAATGTGTGATCAAGCCCGTGAATTTGGTATTAATTCATTTGGGAGATGCGCTG 3511
XX QY 61 GGAATCGACCTATTAATTAATATGATGACATTCGACAGTCTGCTTGATCCAAAAA 120
XX DB 3510 GGAATCGACCTATTAATTAATATGATGACATTCGACAGTCTGCTTGATCCAAAAA 3451
XX QY 121 GGAATCGGACAGAGGATGAACCTCCGCGAATTTTGAAGTGAAGAAATCCTTCGTTG 180
XX DB 3450 GGAATCGGACAGAGGATGAACCTCCGCGAATTTTGAAGTGAAGAAATCCTTCGTTG 3391
XX QY 181 TAAAGGAAGTTTGTGCTTGCGAAGAAACGCGAGATCATCTCTTAACATGAGG 240
XX DB 3390 TAAAGGAAGTTTGTGCTTGCGAAGAAACGCGAGATCATCTCTGAACGAAACAGCA 3331
XX QY 241 AGGAAAAACATGAA 255
```

D8	3330 GTTATAAGCATGAA 3316
RESULT 10	
ID AAL48883/c	
AAL48883 standard; DNA; 6886 BP.	
AC AAL48883;	
XX	
XX 24-OCT-2002 (first entry)	
DE B subtilis panBCD operon vector PAN624.	
XX	
XX 3-(2-hydroxy-3-methyl-butylamino)-propionic acid; HMBPA; alpha-HIV;	
KM alpha-hydroxyisovalerate; pantothenate biosynthesis;	
KW cardiovascular disease; HMGR CoA reductase; hypercholesterolemia;	
XX atherosclerosis; skin aging; PanB; PanE; PanF; PanG; vector; ds.	
XX Bacillus subtilis.	
OS Synthetic.	
PN WO200257476-A2.	
PD 25-JUL-2002.	
PF 19-JAN-2002; 2002WO-US001887.	
PR 19-JAN-2001; 2001US-0263053P.	
PA (OMNI-) OMNIGENE BIOPRODUCTS INC.	
P1 Hermann T, Patterson TA, Pero JG, Yocum RR, Baldenius K, Beck C;	
DR WPI; 2002-608383/65.	
XX Preparation of 3-(2-hydroxy-3-methyl-butylamino)-propionic acid useful	
PT for synthesizing hydroxy methyl glutarate CoA reductase inhibitors,	
PT involves culturing a microorganism under suitable conditions and	
PS detecting or isolating product.	
XX Example 3; Page 68-70; 80pp; English.	
XX The present invention relates to a method of producing 3-(2-hydroxy-3-	
CC methyl-butylamino)-propionic acid (HMBPA), which involves culturing a	
CC microorganism under conditions such that HMBPA is produced, and detecting	
CC or isolating the HMBPA produced by the microorganism. These organisms may	
CC have modified pantothenate biosynthesis enzymes. The method can be used	
CC to produce HMBPA, which can then be used to synthesize inhibitors of	
CC hydroxy methyl glutarate (HMGR) CoA reductase, which are useful for	
CC treating hypercholesterolemia, coronary atherosclerosis progression, and	
CC also produced, which is useful in the prevention of aging of skin, and to	
CC treat skin disorders such as age spots, skin lines, wrinkles, photoaging	
CC and aging. The present sequence is a vector used in the exemplification	
CC of the invention	
SQ Sequence 6886 BP; 1855 A; 1663 C; 1517 G; 1851 T; 0 U; 0 Other;	
Query Match 43.8%; Score 216.2; DB 6; Length 6886;	
Best Local Similarity 98.6%; Pred. No. 3.6e-56;	
Matches 218; Conservative 0; Mismatches 3; Indels 0; Gaps 0.	
OY 1 GGTATGTTGGTCAAGCCCGTTGATTGGTATCTTCATTCAGGAGATGCGCTG 60	
DB 2909 GCTAATGTTGGTCAAGCCCGTTGATTGGTATCTTCATTCAGGAGATGCGCTG 2855	
OY 61 CGAATCGCACCTATTATTAATAATAGATAGACATTGCAGCAGTGCTGCCTTGATCAA AAAA 120	
DB 2849 CGAATCGCACCTATTATTAATAATAGATAGACATTGCAGCAGTGCTGCCTTGATCAA AAAA 2790	
OY 121 GGACTGGGACAAGAGGANTAACTGCCGCACTTTTAGAAGGAAGAATCCCTTCGGTTG 180	
DB 2789 GGACTGGGACAAGAGGANTAACTGCCGCACTTTTAGAAGGAAGAATCCCTTCGGTTG 2730	

Qy	181	TAA	CGA	AGG	TTTTT	GGCTT	GC	TG	CA	GA	AA	CG	CA	GA	T	CG	CA	AT	C	221
Db	2729	TAA	CGA	AGG	TTTTT <td>TGGCTT<td>GC</td><td>TG<td>CA<td>GA<td>AA<td>CG<td>CA<td>GA</td><td>T<td>CG<td>CA</td><td>AT<td>C</td><td>2699</td></td></td></td></td></td></td></td></td></td></td>	TGGCTT <td>GC</td> <td>TG<td>CA<td>GA<td>AA<td>CG<td>CA<td>GA</td><td>T<td>CG<td>CA</td><td>AT<td>C</td><td>2699</td></td></td></td></td></td></td></td></td></td>	GC	TG <td>CA<td>GA<td>AA<td>CG<td>CA<td>GA</td><td>T<td>CG<td>CA</td><td>AT<td>C</td><td>2699</td></td></td></td></td></td></td></td></td>	CA <td>GA<td>AA<td>CG<td>CA<td>GA</td><td>T<td>CG<td>CA</td><td>AT<td>C</td><td>2699</td></td></td></td></td></td></td></td>	GA <td>AA<td>CG<td>CA<td>GA</td><td>T<td>CG<td>CA</td><td>AT<td>C</td><td>2699</td></td></td></td></td></td></td>	AA <td>CG<td>CA<td>GA</td><td>T<td>CG<td>CA</td><td>AT<td>C</td><td>2699</td></td></td></td></td></td>	CG <td>CA<td>GA</td><td>T<td>CG<td>CA</td><td>AT<td>C</td><td>2699</td></td></td></td></td>	CA <td>GA</td> <td>T<td>CG<td>CA</td><td>AT<td>C</td><td>2699</td></td></td></td>	GA	T <td>CG<td>CA</td><td>AT<td>C</td><td>2699</td></td></td>	CG <td>CA</td> <td>AT<td>C</td><td>2699</td></td>	CA	AT <td>C</td> <td>2699</td>	C	2699
RESULT 11																				
ID	AA141227/c																			
AA141227	standard; DNA; 6886 BP.																			
AC	AA141227;																			
XX																				
DT	30-OCT-2002 (first entry)																			
XX																				
DE	Pantothenate related vector PAN624 SEQ ID NO 20.																			
XX																				
XX	3-(2-hydroxy-3-methyl-butrylamino)-propionic acid; HMBPA; Panb; Panb1;																			
XX	pantothenate composition; deregulated pantothenate biosynthetic pathway;																			
XX	deregulated isoleucine-valine; ilv biosynthetic pathway; Panb2; IlvC; ds.																			
XX																				
OS	Unidentified.																			
XX																				
XX	WO200257474-A2.																			
XX																				
XX	25-JUL-2002.																			
XX																				
XX	19-JAN-2002; 2002WO-US001842.																			
XX																				
XX	19-JAN-2001; 2001US-0262995P.																			
XX																				
XX	(OMNI-) OMNIGENE BIOPRODUCTS INC.																			
XX																				
XX	Hermann T, Patterson TA, Pero JG, Yocum RR, Baldeus K, Beck C;																			
XX																				
XX	WPI; 2002-636526/68.																			
XX																				
XX	Preparation of 3-(2-hydroxy-3-methyl-butrylamino)-propionic acid-free																			
XX	pantothenate composition by culturing microorganism having deregulated																			
XX	pantothenate biosynthetic pathway and modified pantothenate biosynthetic																			
XX	enzyme activities.																			
XX																				
XX	Example 3; Page 76-78; 93pp; English.																			
XX																				
XX	The invention relates to a method for the preparation of 3-(2-hydroxy-3-																			
XX	methyl-butrylamino)-propionic acid (HMBPA)-free pantothenate composition																			

Db 2729 TAACGAGAGTTTGTGCTTCAGAGAGAAACGGGAATTC 2689

RESULT 12

ABK73715
ID ABK73715 standard; DNA; 834 BP.

XX
AC ABK73715;

XX
13-AUG-2002 (first entry)

XX
Bacillus licheniformis genomic sequence tag (GST) #1006.

XX
Bacillus licheniformis genomic sequence tag; GST;
altered culture condition; environmental stress;

XX
physiological provocation; ds.

XX
Bacillus licheniformis.

XX
WO200229113-A2.

XX
11-APR-2002.

XX
05-OCT-2001; 2001WO-US031437.

XX
06-OCT-2000; 2000US-00680598.

XX
27-MAR-2001; 2001US-0279526P.

XX
(NOVO) NOVOZYMES BIOTECH INC.

XX
(NOVO) NOVOZYMES AS.

XX
Berka R, Clausen IG;

XX
WPI; 2002-416684/44.

XX
Monitoring differential expression of several genes in first Bacillus
cell relative to expression of same genes in one or more second Bacillus
cells, by using substrate containing Bacillus genomic sequenced tag
array.

XX
Claim 4; SEQ ID NO 1006; 200pp; English.

XX
The invention describes a method of monitoring differential expression of
genes in a first Bacillus cell relative to expression of the genes in
other Bacillus cells, comprising hybridising labelled nucleic acid probes
isolated from Bacillus cells to a substrate containing array of Bacillus
genomic sequenced tags (GST), examining the array, and determining
relative gene expression by an observed hybridisation reporter signal of
a spot in the array. The method is useful for measuring the expression of
genes in a first Bacillus cell relative to expression of the same genes
in one or more second Bacillus cells. The method is useful for monitoring
global expression of several genes from a Bacillus cell, discovering new
genes, identifying possible functions of unknown open reading frames and
monitoring gene copy number variation and stability. Monitoring changes
in expression of genes may be used to provide a representation of the way
in which Bacillus cells adapt to changes in culture conditions,
environmental stress or other physiological provocation. Extensive follow
up characterisation is unnecessary, when one spot on an array equals one
gene or one open reading frame, since sequence information is available.
This sequence represents a genomic sequence tag (GST) used in the method
of the invention. Note: The sequence data for this patent did not form
part of the printed specification, but was obtained in electronic format
directly from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX
Sequence 834 BP; 213 A; 186 C; 240 G; 195 T; 0 U; 0 Other;

Query Match 30.6%; Score 151.2; DB 6; Length 834;
Best Local Similarity 76.2%; Pred. No. 2.5e-36;
Matches 186; Conservative 0; Mismatches 58; Indels 0; Gaps 0;

Qy 251 ATGAAACAAACCTGATTTCTAAATAATGAAGAGTCTGAAGAACGATTTGTCATGCTG 310
1 ATGAGACAAACCTGATTTCTAAATAATGAAGAGAGAGAGAGCCGATGCTATGCTG 60

Db

Qy 311 ACGCTATGATTTATCCGGAGCTAAACTTGCTGAACACGAGGAGTTGACATGATTTTA 370
Db 61 ACTGCTTACGATTAATCCGGCGGCAAAAGCTTGCGAACAGGCGCGGCTGATGATCTTT 120

Qy 371 GTCGGTATTCACCTTGGAATGTCGTCCTCGGCTTGATTCAACTGTCGGTGCACAGTT 430
Db 121 GTGGCGATTTCATCGGCAGATGGTTGTACTCGGCTTACTGACCGATCTGTTACGGTT 180

Qy 431 GCGACATGATTCATCTATCAAAAAGCCGTTAAAAGGGGTGCGCGAATACCTTTATTGTG 490
Db 181 GCTATATGATTCACCATACAAAAGCTGTCAAAACGCGGCAAAAAGATAGTTTGTGTG 240

Qy 491 ACAG 494
Db 241 ACGG 244

RESULT 13

ACA21619
ID ACA21619 standard; DNA; 837 BP.

XX
ACA21619;

XX
19-JUN-2003 (first entry)

XX
Prokaryotic essential gene #3276.

XX
Antisense; ds; prokaryotic essential gene; cell proliferation;
drug design; gene.

XX
Bacillus anthracis.

XX
WO200277183-A2.

XX
03-OCT-2002.

XX
21-MAR-2002; 2002WO-US009107.

XX
21-MAR-2001; 2001US-00815242.

XX
06-SEP-2001; 2001US-00948993.

XX
25-OCT-2001; 2001US-0342923P.

XX
08-FEB-2002; 2002US-00072851.

XX
06-MAR-2002; 2002US-0362699P.

XX
(BLIT-) BLITRA PHARM INC.

XX
Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
Wall D, Trawick JD, Carr GJ, Yamamoto R, Foreyth RA, Xu HH;
P-PSDB; ABU17749.

XX
WPI; 2003-029926/02.

XX
New antisense nucleic acids, useful for identifying proteins or screening
for homologous nucleic acids required for cellular proliferation to
isolate candidate molecules for rational drug discovery programs.

XX
Claim 14; SEQ ID NO 9489; 1766pp; English.

XX
The invention relates to an isolated nucleic acid comprising any one of
the 6213 antisense sequences given in the specification where expression
of the nucleic acid inhibits proliferation of a cell. Also included are:
(1) a vector comprising a promoter operably linked to the nucleic acid
encoding a polypeptide whose expression is inhibited by the antisense
nucleic acid; (2) a host cell containing the vector; (3) an isolated
polypeptide or its fragment whose expression is inhibited by the
antisense nucleic acid; (4) an antibody capable of specifically binding
the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
proliferation or the activity of a gene in an operon required for
proliferation; (7) identifying a compound that influences the activity of
the gene product or that has an activity against a biological pathway
required for proliferation, or that inhibits cellular proliferation; (8)
identifying a gene required for cellular proliferation or the biological

pathway in which a proliferation-required gene or its gene product lies or a gene on which the test compound that inhibits proliferation of an organism acts; (9) manufacturing an antibiotic; (10) profiling a compound's activity; (11) a culture comprising strains in which the gene product is overexpressed or underexpressed; (12) determining the extent to which each of the strains is present in a culture or collection of strains; or (13) identifying the target of a compound that inhibits the proliferation of an organism. The antisense nucleic acid acids are useful for identifying proteins or screening for homologous nucleic acids required for cellular proliferation to isolate candidate molecules for rational drug discovery programs, or for screening homologous nucleic acids required for proliferation in cells other than *S. aureus*, *S. typhimurium*, *K. pneumoniae* or *P. aeruginosa*. The present sequence is one of the target prokaryotic essential genes. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at [ftp://wipo.int/pub/published_pcr_sequences](http://wipo.int/pub/published_pcr_sequences)

Sequence 837 BP; 275 A; 135 C; 210 G; 217 T; 0 U; 0 Other;

Query Match 28.1%; Score 139; DB 8; Length 837;

Best Local Similarity 73.3%; Pred. No. 1.5e-32;

Matches 178; Conservative 0; Mismatches 65; Indels 0; Gaps 0;

```
OY 252 TGAACAACTGATTTCTTAAAGAGAGTCAAGACCGATTGTCATGCTGA 311
    |||||
DB 2 TGAACAACTGATTTCTTAAAGAGAGTCAAGAGTCAAGTCAATGCTGA 61
OY 312 CGGCTTATGATTTCTCGGAGCTAACTTGCTGAACAGCGGAGTTGACATGATTTAG 371
    |||||
DB 62 CGGCTTATGATTTCTCGGAGCTAACTTGCTGAACAGCGGAGTTGATGATTTAG 121
OY 372 TCGGATTCATCTTGGAATGCTGCTCGGCTTGAATCAATGCTGCTGACATG 431
    |||||
DB 122 TTGAGATTTCTCTCGGAGTGTGTGCTCGGCTAAGATTCACAGTAACAGTAG 181
OY 432 CGGATGATTCATCATACAAAGCGCTTAAAGGGTGGCCGGAATCCTTATTTGA 491
    |||||
DB 182 AGGATGATTCATCATACAAAGCGCTTAAAGGGTGGCCGGAATCCTTATTTGA 241
OY 492 CAG 494
    |||||
DB 242 CTG 244
```

RESULT 14

ABQ70448 standard; DNA; 1613 BP.

AC ABQ70448;

XX 29-AUG-2003 (revised)

DT 29-AUG-2002 (first entry)

DE *Listeria monocytogenes* 4b contig DNA sequence #390.

XX Antibacterial; *Listeria*; food contamination; mutational analysis;

KW infection; ds.

OS *Listeria monocytogenes* ATCC 19115.

XX WO200228891-A2.

XX 11-APR-2002.

XX 04-OCT-2001; 2001WO-FR003061.

PR 04-OCT-2000; 2000FR-00012697.

PA (INSP) INST PASTEUR.

XX (CNRS) CNRS CENT NAT RECH SCI.

PI Kunst F, Glaeser P;

XX WPI; 2002-332479/37.

PT New genomic sequences from *Listeria* species, useful for detection,

PT treatment and prevention of infection, also related polypeptides,

PS antibodies and modulators.

XX Claim 14; SEQ ID NO 3261; 180bp; French.

The present invention relates to nucleic acid sequences (ABQ67188-ABQ71212) from *Listeria* sp. The sequences are useful as probes and primers for identification and/or detection of *Listeria* (e.g. as contaminants in foods, or mutational analysis) and for analysis of gene expression. Proteins encoded by the nucleic acid sequences can be used to screen for compounds that modulate gene expression, replication and pathogenicity of *Listeria* (potential therapeutic agents), also for treating infections by *Listeria*, and are useful as immunogens in anti-*Listeria* vaccines. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at [ftp://wipo.int/pub/published_pcr_sequences](http://wipo.int/pub/published_pcr_sequences). (Updated on 29-AUG-2003 to standardise OS field)

Sequence 1613 BP; 540 A; 295 C; 366 G; 412 T; 0 U; 0 Other;

Query Match 27.7%; Score 137; DB 6; Length 1613;

Best Local Similarity 70.8%; Pred. No. 8.7e-32;

Matches 182; Conservative 0; Mismatches 75; Indels 0; Gaps 0;

```
OY 238 AGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 297
    |||||
DB 146 AGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 205
OY 298 GATGTCATGTCGACCGCTTATGATTTCCGGAGCTAACTTGCTGAACAGCGGAGT 357
    |||||
DB 206 AATACAGATGATTAACCGGATGATGATTTCTTGCTGAAGATGATGAACAGAGAGC 265
OY 358 TGACATGATTTAGTCGGATGATTCATTTGGAATGTCGCTCGGCTTGAATCACTGT 417
    |||||
DB 266 TGATGATTTTATTCGGGATTCATCTTGAATGATGATTTAGCTACAGTTCACTGT 325
OY 418 CGGTCGACAGTTGCGGACATGATTCATACAAAGCGCTTAAAGGGTGGCCGGA 477
    |||||
DB 326 ACCAGTTACAGTATGATGATTCATACAAAGCGGTCGAAGCGGGGACACGA 385
OY 478 TACCTTATTTGTCAG 494
    |||||
DB 386 TACATTTGTTGTTACAG 402
```

RESULT 15

ABQ69245_20/c Continuation (21 of 31) of ABQ69245 from base 2000001 (*Listeria innocua* DNA sequence #68

WP Sequence split into 31 fragments LOCUS ABQ69245 Accession Abq69245

Fragment Name	Begin	End
WP ABQ69245_00	1	110000
WP ABQ69245_01	100001	210000
WP ABQ69245_02	200001	310000
WP ABQ69245_03	300001	410000
WP ABQ69245_04	400001	510000
WP ABQ69245_05	500001	610000
WP ABQ69245_06	600001	710000
WP ABQ69245_07	700001	810000
WP ABQ69245_08	800001	910000
WP ABQ69245_09	900001	1010000
WP ABQ69245_10	1000001	1110000
WP ABQ69245_11	1100001	1210000
WP ABQ69245_12	1200001	1310000
WP ABQ69245_13	1300001	1410000
WP ABQ69245_14	1400001	1510000
WP ABQ69245_15	1500001	1610000
WP ABQ69245_16	1600001	1710000
WP ABQ69245_17	1700001	1810000
WP ABQ69245_18	1800001	1910000

WP	AB069245_19	1900001	2010000
WP	AB069245_20	2000001	2110000
WP	AB069245_21	2100001	2210000
WP	AB069245_22	2200001	2310000
WP	AB069245_23	2300001	2410000
WP	AB069245_24	2400001	2510000
WP	AB069245_25	2500001	2610000
WP	AB069245_26	2600001	2710000
WP	AB069245_27	2700001	2810000
WP	AB069245_28	2800001	2910000
WP	AB069245_29	2900001	3010000
WP	AB069245_30	3000001	3011208

Query Match 27.7%; Score 137; DB 6; Length 110000;
Best Local Similarity 70.8%; Pred. No. 5.5e-31;
Matches 182; Conservative 0; Mismatches 75; Indels 0; Gaps 0;

QY	238	AGAGAGAGAAAAATGAAAAACAAACCTGATTTCTAAAAATGAAGAGCTGAAGAAC	297
DB	31566	AGAGCTAAGAAAAATGAAAAACAGTGAAGCTTTTGTCTATGAAGAAAAACGAGAAAA	31507
QY	298	GATTGTCATGCTGACCGCTTATGATTATCCGGCAGCTAACTTGCTGAACAAGCGGGAGT	357
DB	31506	AATCAGATGATTACCGCTTACGATTATCTTCTGCTAAGAAATGTAGAACAAAGCAGAAC	31447
QY	358	TGACATGATTTTAATGCGGATTCCTGGAATGGTCCTCGGCTTGATTCACCTGT	417
DB	31446	AGATATGATTTAGTTGATTCCTGGAATGGTATGATTAAGCTTACGATTCACCTGT	31387
QY	418	CGGTGTGACAGTTGCGGACATGATCATACAAAGCCGTTAAAGGGGTGCGCGAA	477
DB	31386	ACCGTAAACATGATGATATTCATCATACGAAGCTGTGAACGAGGGGCGCCAA	31327
QY	478	TACCTTATTTGACAG	494
DB	31326	TACTTTGTGTAAACGG	31310

Search completed: September 8, 2005, 20:33:22
Job time : 7325 secs

27	41.8	8.5	9454	4	US-09-949-016-123398	Sequence 123398, App
26	43.4	8.8	618	4	US-09-710-279-1827	Sequence 1827, App
25	43.8	8.9	755	3	US-08-998-416-570	Sequence 570, App
24	45.8	9.3	4411529	3	US-09-103-8408-1	Sequence 1, Appl
23	45.8	9.3	4403765	3	US-09-103-8408-2	Sequence 2, Appl
22	47.4	9.6	100848	4	US-09-536-002-39	Sequence 39, Appl
21	47.4	9.6	2164	3	US-09-318-7934-1	Sequence 1, Appl
20	47.4	9.6	2164	3	US-09-318-7944-1	Sequence 5, Appl
19	47.4	9.6	855	4	US-09-540-236-5	Sequence 5, Appl
18	49.2	10.0	813	4	US-09-489-039A-5090	Sequence 5090, App
17	50.8	10.3	846	4	US-09-553-681A-2061	Sequence 2061, App
16	52	10.5	2538	3	US-09-221-017B-369	Sequence 369, App
15	54.2	11.0	2364	4	US-09-790-988-1	Sequence 1, Appl
14	54.2	11.0	2364	4	US-09-902-540-506	Sequence 506, App
13	54.2	11.0	889	4	US-09-902-540-6623	Sequence 6623, App
12	59.2	12.0	1005	4	US-09-252-991A-10798	Sequence 10798, A
11	59.2	12.0	912	4	US-09-252-991A-10351	Sequence 10351, A
10	61.2	12.4	9381	3	US-09-453-702B-7	Sequence 7, Appl
9	74.2	15.0	855	4	US-09-338-352-3820	Sequence 3820, App
8	78	15.8	3301	4	US-09-710-279-4209	Sequence 4209, App
7	87.4	17.7	828	4	US-09-134-000C-3123	Sequence 3123, App
6	87.4	17.7	828	4	US-09-107-532A-3559	Sequence 3559, App
5	99.6	20.2	4549	4	US-08-781-986A-174	Sequence 174, App
4	99.6	20.2	4549	4	US-08-956-171E-174	Sequence 174, App
3	103.2	20.9	792	4	US-09-710-279-1417	Sequence 1417, App
2	111.6	22.6	825	3	US-09-134-001C-1569	Sequence 1569, App
1	117.2	23.7	1880	4	US-09-710-279-4426	Sequence 4426, App

28	41.8	8.5	9455	4	US-09-949-016-14650	Sequence 14650, A
29	41.6	8.4	560	4	US-09-903-814A-1	Sequence 1, Appli
30	41.6	8.4	560	4	US-10-723-061-1	Sequence 1, Appli
31	39.2	7.9	87863	4	US-09-949-016-14402	Sequence 14402, A
32	38.6	7.8	7248	1	US-08-232-466-14	Sequence 14, Appli
33	37.2	7.5	11652	1	US-09-818-512-3	Sequence 17611, A
34	36.8	7.4	601	4	US-09-949-016-112617	Sequence 16625, A
35	36.8	7.4	39289	4	US-09-949-016-15625	Sequence 15845, A
36	36	7.3	133559	4	US-09-949-016-15845	Sequence 15602, A
37	35.4	7.2	134890	4	US-09-949-016-15602	Sequence 14870, A
38	35.4	7.1	601	4	US-09-949-016-94870	Sequence 14870, A
39	35.2	7.1	30032	4	US-09-949-016-13933	Sequence 13933, A
40	35.2	7.1	88366	4	US-09-949-016-14502	Sequence 14502, A
41	35.2	7.1	58108	4	US-09-949-016-13383	Sequence 13383, A
42	34.5	7.1	10965	4	US-09-949-016-14097	Sequence 14097, A
43	34.8	7.0	601	4	US-09-949-016-15464	Sequence 15464, A
44	34.8	7.0	601	4	US-09-949-016-182713	Sequence 182713, A
45	34.8	7.0	67911	4	US-09-949-016-16779	Sequence 16799, A

ALIGNMENTS

```

RESULT 1
US-09-710-279-4426/C
; Sequence 4426, Application US/09710279
; Patent No. 6703492
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: P13480US
; CURRENT APPLICATION NUMBER: US/09/710,279
; CURRENT FILING DATE: 2000-11-09
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4426
; LENGTH: 1880
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: nucleic acid sequence
US-09-710-279-4426

```

Query Match	23.7%	Score 117.2;	DB 4;	Length 1880;
Best Local Similarity	65.9%	Pred. No. 1.5e-27;		
Matches 170; Conservative	0;	Mismatches 88;	Indels 0;	Gaps 0;

OY	232	AAATGATGGAGGAGGAAAACATGAAAAACAACCTGGATTTTCTAAAAAAGAGAGTCTGA	253
Db	967	AAATGAGGGTGAATGAAATGAAAACCTTAAATCATTTAAACAAAATGAAGGATCACA	908
OY	292	AGAAACGATTTGTCATCTGACCCGCTTAGATTAATCCGCACTAAACTCTGAAACAAGC	351
Db	907	GCAAAAGATTTCTATGTTACAGCTTAAGATTTATCTTAGCTAAGCAAGCAACAAGC	848
OY	352	GGAGATGACATGATTTTATGTCGGTGATTCACCTTGAATGGTCGTCTCGGCTCTTGATTC	411
Db	847	TGAATATGACATATATTTGTAAGAGATTTCTTAGGAATGACAGTGTAGGATATGATAG	788
OY	412	AACTGTCGATGACAGTTCGGACACATGATTCATCATACAAAAGCCGTTAAAAAGGGTGC	471
Db	787	TACGTCAAGTTACATTAACGATGATGATTCATCATGATGTAAGCGCTGTTAAAAAGGGTGC	728
OY	472	GCCGATACCTTTATTTGT	489
Db	727	TTGAGATCAATTTATAGT	710

RESULT 2
US-09-134-001C-1569

; Sequence 1569, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 1569
; LENGTH: 825
; TYPE: DNA
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-1569

Query Match 22.6%; Score 111.6; DB 3; Length 825;
Best Local Similarity 66.8%; Pred. No. 6.4e-26;
Matches 159; Conservative 0; Mismatches 79; Indels 0; Gaps 0;

Qy 252 TGAACCAAACTGGATTTCTTAAATGAAGAGCTTGAAGACCGATTGTCATGCTGA 311
Db 8 TGAACCACTTAAATCATTTAAACAAATGAAGAGCTACACGAAAGATTTCTATGTTA 67
Qy 312 CCGCTTATGATTTATCCGCGAGCTAAATCTGCTGAACACGGGAGTTGACATGATTTAG 371
Db 68 CAGCTTATGATTTATCTAGTGTGAAGCAACCAAGCTAAATGACATGATTTAG 127
Qy 372 TCGGATGATTCACCTTGAATGCTGCTCGGCTTGAATTCATGCTGCTGACAGTTG 431
Db 128 TAGGAGATTTCTTGAAGATGACAGTGTGATGATGATGATGATGATGATGATGATG 187
Qy 432 CGGACATGATTCATCATCAAAAGCGCTTAAAGGGGTGCGCGGATACCTTTATGT 489
Db 188 ACGATATGATTCATCATGTTAAGGCTGTTAAAGGGGTGCTTCAATGATTTATAGT 245

RESULT 3
US-09-710-279-1417
; Sequence 1417, Application US/09710279
; Patent No. 6703492
; GENERAL INFORMATION:
; APPLICANT: KIMBERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PU3480US
; CURRENT APPLICATION NUMBER: US/09/710,279
; CURRENT FILING DATE: 2000-11-09
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1417
; LENGTH: 792
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: nucleic acid sequence
US-09-710-279-1417

Query Match 20.9%; Score 103.2; DB 4; Length 792;
Best Local Similarity 67.9%; Pred. No. 3.3e-23;
Matches 144; Conservative 0; Mismatches 68; Indels 0; Gaps 0;

Qy 278 ATGAAGAGTCTGAAGAACCGATTCATGCTGACCGCTTATGATTTCCGCGAGCTAA 337
Db 1 ATGAAGCATCACACCAAAAGTTTCTATGCTTACAGCTTATGATTTATCTTAGTCTAAG 60
Qy 338 CTGCTGAACAAGCGGAGGTGACATGATTTAGTCGGTATTCATTGGAATGCTGCTC 397

Db 61 CAAGCAACAAGCTGAATTTGACATGATTTTGTAGAGATTTCTTAGAAGACAGTG 120
Qy 398 CTGGCCCTGATTTCACTGTCGGTGTGACAGTTGCGGACATGATTCATCAACAAAGCC 457
Db 121 TTAGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 180
Qy 458 GTTAAAGAGGTGCGCGCAATACCTTTATGT 489
Db 181 GTTAAAGAGGTGCTTCAGATACATTTATAGT 212

RESULT 4
US-08-956-171E-174/C
; Sequence 174, Application US/08956171E
; Patent No. 6593114
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; GIL H. Choi
; Patrick S. Dillon
; Craig A. Rosen
; Steven C. Barash
; Michael R. Pannon
; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 5256
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/956,171E
; FILING DATE: 20-Oct-1997
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/009,861
; FILING DATE: January 5, 1996
; APPLICATION NUMBER: 08/781,986
; FILING DATE: January 3, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Mark J. Hyman
; REGISTRATION NUMBER: 46,789
; REFERENCE/DOCKET NUMBER: PB248P1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (240) 314-1224
; TELEFAX: (301) 309-8439
; INFORMATION FOR SEQ ID NO: 174:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4549 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 174:
US-08-956-171E-174

Query Match 20.2%; Score 99.6; DB 4; Length 4549;
Best Local Similarity 61.6%; Pred. No. 1.2e-21;
Matches 159; Conservative 0; Mismatches 99; Indels 0; Gaps 0;

Qy 232 AACATGAGAGAGGAGAAAACATGAACAAACTGATTTTCTAATAATGAAGAGTCTGA 291
Db 4535 AAAAAGAGGTGAATTAACATTTGAACAGTAAGTCATTAATGATATGAACAAAGCA 4476
Qy 292 AGAACCATTTGCTGCTGACCGCTTATGATTTATCCGCGAGCTTAAATTTGCTGAACAAG 351
Db 4475 AACTAAATTTCTATGATTAACGCTTATGATTTTCCAAAGTGTAAACAAAGTGAAGCAGC 4416


```
Oy 374 GGTGATTCATTTGGAATGCTGCTCGGCTTGATTCAGTGTGTCAGATTGCG 433
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 127 GGAGATTCCTTTGGATGTGATGTCTAGAGACATGAAATACCTTCAATGAAATGAA 186
Oy 434 GACATGATTCATCATACAAAGCCGTTAAAGGGGTGGCCGGAATACCTTATTTGACA 493
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 187 GACATGATTCATCATACATAGACAGCTTACAGAGAGAGCTAAGAACCTTTAGTTGTGCA 246
Oy 494 G 494
Db 247 G 247
```

RESULT 7
US-09-134-000C-3123
Sequence 3123, Application US/09134000C

```
Patent No. 6617156
GENERAL INFORMATION:
APPLICANT: Lynn Doucette-Stamm et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
TITLE OF INVENTION: ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: 032796-032
CURRENT APPLICATION NUMBER: US/09/134,000C
PRIOR FILING DATE: 1998-08-13
PRIOR APPLICATION NUMBER: US 60/055,778
NUMBER OF SEQ ID NOS: 6812
SOFTWARE: PatentIn version 3.1
SEQ ID NO 3123
LENGTH: 828
TYPE: DNA
ORGANISM: Enterococcus faecalis
US-09-134-000C-3123
```

Query Match 17.7%; Score 87.4; DB 4; Length 828;
Best Local Similarity 60.2%; Pred. No. 4,7e-18;
Matches 145; Conservative 0; Mismatches 96; Indels 0; Gaps 0;

```
Oy 254 AAAACAAAACGATTTCTTAAATGAAGAGCTGGAAGAACGATTTGTCAGTGCAC 313
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 7 AATACAGAGTTACATTTAAAGATCTTAAATTAAGAAATGAATAATTAACATGCTTACA 66
Oy 314 GCTTATGATTTATCCGAGCTAAACCTTCTGTAACAGCGGAGTTGACATGATTTTATGTC 373
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 67 GCTTATGATTTATTAACAAGCAAAATAATATGATGAAGCTGGGATTAATGAAATATTAGTA 126
Oy 374 GGTGATTCATTTGGAATGCTGCTCGGCTTGATTTCACTGTGCGTGTGACAGTTGCG 433
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 127 GGAGATTCCTTTGGATGTGATGTCTAGACATGAAATACATTTCACTAATCATGGA 186
Oy 434 GACATGATTCATCATACAAAGCCGTTAAAGGGGTGGCCGGAATACCTTATTTGACA 493
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 187 GACATGATTCATCATACATAGACAGCTTACAGAGAGAGCTAAGAACCTTTAGTTGTGCA 246
Oy 494 G 494
Db 247 G 247
```

RESULT 8
US-09-710-279-4209/c
Sequence 4209, Application US/09710279
Patent No. 6703492
GENERAL INFORMATION:
APPLICANT: KIMMERLY, WILLIAM JOHN
TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
FILE REFERENCE: PUS480US
CURRENT APPLICATION NUMBER: US/09/710,279
PRIOR FILING DATE: 2000-11-09
PRIOR APPLICATION NUMBER: 60/164,258
NUMBER OF SEQ ID NOS: 4472
SOFTWARE: PatentIn Ver. 2.1

```
; SEQ ID NO 4209
; LENGTH: 3301
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
US-09-710-279-4209
```

Query Match 15.8%; Score 78; DB 4; Length 3301;
Best Local Similarity 68.4%; Pred. NO. 1.1e-14;
Matches 108; Conservative 0; Mismatches 50; Indels 0; Gaps 0;

```
Oy 332 GCTAACTGCTGAACAAGCGGAGTTGACATGATTTTATGCTGATTCATCTGAAATG 391
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3297 GCTAAGAGACACACAGACAGCTGAATTTGATGATATTTTGTAGAGATTTCTTAGAGATG 3238
Oy 392 GTGCTCTCGGCTTATTTCACTGTCGTTGTGACAGTTGCGGACATGATTCATACACA 451
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3237 ACAGTGTAGGATATGATAGTACTGTTCAAGTTACATTTGAACATATGATTCATATGAT 3178
Oy 452 AAAGCCCTTAAAGGGGTGGCCGAATACCTTATTTGT 489
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3177 AAGGCTGTAAAGAGGTGCTTCAGATACATTTATAGT 3140
```

RESULT 9
US-09-328-352-3820
Sequence 3820, Application US/09328352
Patent No. 6562958

```
GENERAL INFORMATION:
APPLICANT: Gary L. Breton et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: GTC99-03PA
CURRENT APPLICATION NUMBER: US/09/328,352
PRIOR FILING DATE: 1999-06-04
NUMBER OF SEQ ID NOS: 8252
SEQ ID NO 3820
LENGTH: 855
TYPE: DNA
ORGANISM: Acinetobacter baumannii
US-09-328-352-3820
```

Query Match 15.0%; Score 74.2; DB 4; Length 855;
Best Local Similarity 56.3%; Pred. No. 9.3e-14;
Matches 139; Conservative 0; Mismatches 108; Indels 0; Gaps 0;

```
Oy 248 AACATGAAAACAAACGATTTTCTTAAATGAAGAGCTGGAAGAACGATTTGTCATG 307
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 43 ATCATGATTTAGTCTAAGTGACTTAAGAAATTTAAAGCCGAGAGCGTAAGTTCTTTGC 102
Oy 308 CTGACCGCTTATGATTTATCCGAGCTAAACCTTGCTGAACAAGCGGAGTTGACATATT 367
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 103 CTAACTTTTACGATGCAATGATGCAAAAGCAATGAACTTGTGAAATTTGATACATC 162
Oy 368 TTATGCGTATTCATTTGGAATGATGCTGCTCGGCTTATTTCAATCACTGCGGTGACA 427
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 163 TTATTTGATGATTTCTTTGGAATGCAATTCAGGGGCGTACATCACTTAACTTAACT 222
Oy 428 GTTGGGACATGATTCATACAAAGCCGTTAAAGGGGTGGCCGGAATACCTTTATTT 487
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 223 GTTGAACATAGGCTTATCATACAGCGGAGTCCGTGCGGTATACGATGCTTTAATT 282
Oy 488 GTGACAG 494
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 283 ATGACTG 289
```

RESULT 10
US-09-453-702B-7/c
Sequence 7, Application US/09453702B
Patent No. 6365723

GENERAL INFORMATION:
APPLICANT: Blattner, Frederick R.
Burland, Valerie
Perna, Nicole T.
Plunkett, Guy
Welch, Rod
TITLE OF INVENTION: No. 6365723e1 Sequences of E. coli 0157
NUMBER OF SEQUENCES: 265
CORRESPONDENCE ADDRESS:
ADDRESSEE: Quarles & Brady
STREET: 1 South Pinckney Street
CITY: Madison
STATE: WI
COUNTRY: US
ZIP: 53701-2113
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch. 1.44Mb storage
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Word Perfect 8.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/453,702B
FILING DATE: 03-Dec-1999
CLASSIFICATION: <unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/110,955
FILING DATE: 04-DEC-1998
ATTORNEY/AGENT INFORMATION:
NAME: Seay, Nicholas J.
REGISTRATION NUMBER: 27386
REFERENCE/DOCKET NUMBER: 960296.95017
TELECOMMUNICATION INFORMATION:
TELEPHONE: (608) 251-5000
TELEFAX: (608) 251-9166
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 9381
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
SEQUENCE DESCRIPTION: SEQ ID NO: 7:
US-09-453-702B-7

Query Match 12.4%; Score 61.2; DB 3; Length 9381;
Best Local Similarity 54.4%; Pred. No. 5.5e-09;
Matches 123; Conservative 0; Mismatches 103; Indels 0; Gaps 0;

QY 269 TTTCTAAATGAGAGCTGAGAGACCGATTGTCATGCTGACCGCTTATGATTATCCG 328
DB 1935 TTGCAAAATGTAACAGAAAACGCTTCGCCACATCACCGCTACGACTACAGC 1876

QY 329 GCACTAAACTGCTGAACACCGGAGTTCATGATTTTATGCGGTGATTCACCTTGA 388
DB 1875 TTCGAAAACGTTTCCGAGAGAGCGCTGAACGTAATGCTGTGGCGCACTGCGTGGG 1816

QY 389 ATGTCGTCCTCGGCTTGAATCACTGTCGCTGTGACAGTTGCGGACATGATTCATCAT 448
DB 1815 ATGACCGTTACGGGCGCATTAATCTACCTTCCCGTTACCGTTGAAGATATGCCCTACAC 1756

QY 449 ACAAAAGCGTTAAAGGGGTGCGCGGAATACCTTTATTTGACAG 494
DB 1755 ACTACCGAGTACGTGCTGCGCGCAAACTGCTGCTGTGGCTG 1710

RESULT 11
US-09-252-991A-10351
Sequence 10351, Application US/09252991A
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS

FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 10351
LENGTH: 912
TYPE: DNA
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-10351

Query Match 12.0%; Score 59.2; DB 4; Length 912;
Best Local Similarity 54.6%; Pred. No. 7.2e-09;
Matches 118; Conservative 0; Mismatches 98; Indels 0; Gaps 0;

QY 279 TGAAGAGTCTGAAGACCGATTGTCATGCTGACCGCTTATGATTATCCGACCTAAAC 338
DB 146 TGAAGCAGAGCGCGGAGAAATGCCATGCTGACTGACGACGACCACTTGCCACA 205

QY 339 TTGCTGAACAGCGGAGATTGACATGATTTTATGCTGATTCATTGGAATGCTGTC 398
DB 206 CCGCAGCGAGCGCGGCTGACGCTTCTGTGCGAAGATTCGCTCGAATGATTC 265

QY 399 TCGGCTTGAATCACTGTCGTGTGACAGTTGGGACATGATCATCATCAAAAGCG 458
DB 266 AGGGTCACAGACGACCTCGCGGTGACGAAAGAAATGCTATTCACAGCGCTGCG 325

QY 459 TTAAGGGGTGCGCGCAATACCTTTATTTGACAG 494
DB 326 TGAACGGGTAAACAAAGTTCCTGATGTCACCG 361

RESULT 12
US-09-252-991A-10798/C
Sequence 10798, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 10798
LENGTH: 1005
TYPE: DNA
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-10798

Query Match 12.0%; Score 59.2; DB 4; Length 1005;
Best Local Similarity 54.6%; Pred. No. 7.6e-09;
Matches 118; Conservative 0; Mismatches 98; Indels 0; Gaps 0;

QY 279 TGAAGAGTCTGAAGACCGATTGTCATGCTGACCGCTTATGATTATCCGACGCTAAAC 338
DB 773 TGAAGCAGAGCGCGGAGAAATGCCATGCTGACTGACGACGACCACTTGCCACA 714

QY 339 TTGCTGAACAGCGGAGATTGACATGATTTTATGCTGATTCATTGGAATGCTGTC 398
DB 713 CCGCAGCGAGCGCGGCTGACGCTTCTGTGCGAAGATTCGCTCGAATGATTC 654

QY 399 TCGGCTTGAATCACTGTCGTGTGACAGTTGGGACATGATCATCATCAAAAGCG 458
DB 653 AGGGTCACAGACGACCTCGCGGTGACGAAAGAAATGCTATTCACAGCGCTGCG 594

Best Local Similarity 54.9%; Pred. No. 5e-07;
Matches 107; Conservative 0; Mismatches 88; Indels 0; Gaps 0

Job time : 160 secs

Search completed: September 8, 2005, 22:08:13
Job time : 160 secs

Query Match	98.2%	Score 485	DB 21	Length 2363
Best Local Similarity	100.0%	Pred. No. 5.8e-118		
Matches	485	Conservative 0	Mismatches 0	Indels 0
				Gaps 0
QY	10	TTGGTACAAAGCCCGTTGATTTTGGTATCTCCATTTGGGACGATATCGCTCGCACTGCA	69	
Db	1	TTGGTACAAAGCCCGTTGATTTTGGTATCTCCATTTGGGACGATATCGCTCGCACTGCA	60	
QY	70	CCTATTATTAATAATGATAGACATTTGACAGACTGCTGCTTGAATCCAAAAMAAAGACTGGGA	129	
Db	61	CCTATTATTAATAATGATAGACATTTGACAGACTGCTGCTTGAATCCAAAAMAAAGACTGGGA	120	
QY	130	CAGAGGGATGAATCTGCCGAACTTTAGAAAGTGAAGATCTTCTCGTTGTAACGGAAG	189	
Db	121	CAGAGGGATGAATCTGCCGAACTTTAGAAAGTGAAGATCTTCTCGTTGTAACGGAAG	180	
QY	190	GTTTTTTGGCTTGCAGAAAGAAACGGCAGATCATCTCTCTTAAACATGAGAGAGAGAAA	249	
Db	181	GTTTTTTGGCTTGCAGAAAGAAACGGCAGATCATCTCTCTTAAACATGAGAGAGAGAAA	240	
QY	250	CATGAAAACAAAACGTGATTTTCTTAAATGAAAGAGCTGAAAGAACGATGTCAATGCT	309	
Db	241	CATGAAAACAAAACGTGATTTTCTTAAATGAAAGAGCTGAAAGAACGATGTCAATGCT	300	
QY	310	GACCGCTTATGATTAATCCGGCAGCTTAACTTGCTGACAAAGCGGAGGTGAATGATATTT	369	
Db	301	GACCGCTTATGATTAATCCGGCAGCTTAACTTGCTGACAAAGCGGAGGTGAATGATATTT	360	
QY	370	AGTCGGTATTCACCTTGGAAATGTCCTCTCGGCTTGATTCACCTGTGCTGTGACAGT	429	

```

OY      430  TGGGACATGATTCATCATACAAAAGCCGTTAAAGGGGGTGGCCGAATACCTTATGT 480
        |||||
Db      421  TGGGACATGATTCATCATACAAAAGCCGTTAAAGGGGGTGGCCGAATACCTTATGT 480

OY      490  GACAG 494
        |||||
Db      481  GACAG 485

RESULT 3
US-10-984-449-94/c
; Sequence 94, Application US/10984449
; Publication No. US20050089973A1
; GENERAL INFORMATION:
; APPLICANT: Yocum, R. Rogers
; APPLICANT: Patterson, Thomas A.
; APPLICANT: Hermann, Theron
; APPLICANT: Petro, Janice G.
; TITLE OF INVENTION: METHODS AND MICROORGANISMS FOR PRODUCTION OF
; TITLE OF INVENTION: PANTO-COMPOUNDS
; FILE REFERENCE: BGI-141CPN
; CURRENT APPLICATION NUMBER: US/10/984,449
; CURRENT FILING DATE: 2004-11-08
; PRIOR APPLICATION NUMBER: USN 09/667,569
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: USN 09/400,494
; PRIOR FILING DATE: 1999-09-21
; PRIOR APPLICATION NUMBER: USN 60/210,072
; PRIOR FILING DATE: 2000-06-07
; PRIOR APPLICATION NUMBER: USN 60/221,836
; PRIOR FILING DATE: 2000-07-28
; PRIOR APPLICATION NUMBER: USN 60/227,860
; PRIOR FILING DATE: 2000-08-24
; NUMBER OF SEQ ID NOS: 94
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 94
; LENGTH: 7381
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:

```


PRIOR FILING DATE: 2000-08-24
NUMBER OF SEQ ID NOS: 94
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 93
LENGTH: 8503
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Recombinant
US-10-984-449-93

Query Match 52.0%; Score 257; DB 21; Length 8503;
Best Local Similarity 100.0%; Pred. No. 1,36-67;
Matches 257; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 238 AGAGAGAGAAAAATGAAAAAAGAACTGATTTTCTAAAAATGAAAGAGTCTGAAGAAC 297
DB 3656 AGAGAGAGAAAAATGAAAAAAGAACTGATTTTCTAAAAATGAAAGAGTCTGAAGAAC 3597
QY 298 GATTGTCATGCTGACCGCTTATGATTATCCGAGAGCTAACTGCTGAAGAGGGGAGT 357
DB 3596 GATTGTCATGCTGACCGCTTATGATTATCCGAGAGCTAACTGCTGAAGAGGGGAGT 3537
QY 358 TGACATGATTTTATGCTGCTGATTCACCTGGAATGCTGCTCGGCTTGAATCACTGT 417
DB 3536 TGACATGATTTTATGCTGCTGATTCACCTGGAATGCTGCTCGGCTTGAATCACTGT 3477
QY 418 CGGTGTGACAGTTCGGGACATGATCCATCAATCAAAAGCCGTTAAAGGGGTGCGCGAA 477
DB 3476 CGGTGTGACAGTTCGGGACATGATCCATCAATCAAAAGCCGTTAAAGGGGTGCGCGAA 3417
QY 478 TACCTTATTGTGACAG 494
DB 3416 TACCTTATTGTGACAG 3400

RESULT 7

US-10-984-449-76/c

Sequence 76, Application US/10984449
Publication No. US2005008973A1
GENERAL INFORMATION:
APPLICANT: Yocum, R. Rogers
APPLICANT: Patterson, Thomas A.
APPLICANT: Hermann, Theron
APPLICANT: Pero, Janice G.
TITLE OF INVENTION: METHODS AND MICROORGANISMS FOR PRODUCTION OF
FILE REFERENCE: BGI-141CPCN
CURRENT APPLICATION NUMBER: US/10/984,449
CURRENT FILING DATE: 2004-11-08
PRIOR APPLICATION NUMBER: USSN 09/667,569
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: USSN 09/400,494
PRIOR FILING DATE: 1999-09-21
PRIOR APPLICATION NUMBER: USSN 60/210,072
PRIOR FILING DATE: 2000-06-07
PRIOR APPLICATION NUMBER: USSN 60/221,836
PRIOR FILING DATE: 2000-07-28
PRIOR APPLICATION NUMBER: USSN 60/227,860
PRIOR FILING DATE: 2000-08-24
NUMBER OF SEQ ID NOS: 94
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 76
LENGTH: 10801
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Recombinant
US-10-984-449-76

Query Match 52.0%; Score 257; DB 21; Length 10801;

Best Local Similarity 100.0%; Pred. No. 1,56-67;
Matches 257; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 238 AGAGAGAGAAAAATGAAAAAAGAACTGATTTTCTAAAAATGAAAGAGTCTGAAGAAC 297
DB 3656 AGAGAGAGAAAAATGAAAAAAGAACTGATTTTCTAAAAATGAAAGAGTCTGAAGAAC 3597
QY 298 GATTGTCATGCTGACCGCTTATGATTATCCGAGAGCTAACTGCTGAAGAGGGGAGT 357
DB 3596 GATTGTCATGCTGACCGCTTATGATTATCCGAGAGCTAACTGCTGAAGAGGGGAGT 3537
QY 358 TGACATGATTTTATGCTGCTGATTCACCTGGAATGCTGCTCGGCTTGAATCACTGT 417
DB 3536 TGACATGATTTTATGCTGCTGATTCACCTGGAATGCTGCTCGGCTTGAATCACTGT 3477
QY 418 CGGTGTGACAGTTCGGGACATGATCCATCAATCAAAAGCCGTTAAAGGGGTGCGCGAA 477
DB 3476 CGGTGTGACAGTTCGGGACATGATCCATCAATCAAAAGCCGTTAAAGGGGTGCGCGAA 3417
QY 478 TACCTTATTGTGACAG 494
DB 3416 TACCTTATTGTGACAG 3400

RESULT 8

US-10-984-449-23

Sequence 23, Application US/10984449
Publication No. US2005008973A1
GENERAL INFORMATION:
APPLICANT: Yocum, R. Rogers
APPLICANT: Patterson, Thomas A.
APPLICANT: Hermann, Theron
APPLICANT: Pero, Janice G.
TITLE OF INVENTION: METHODS AND MICROORGANISMS FOR PRODUCTION OF
FILE REFERENCE: BGI-141CPCN
CURRENT APPLICATION NUMBER: US/10/984,449
CURRENT FILING DATE: 2004-11-08
PRIOR APPLICATION NUMBER: USSN 09/667,569
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: USSN 09/400,494
PRIOR FILING DATE: 1999-09-21
PRIOR APPLICATION NUMBER: USSN 60/210,072
PRIOR FILING DATE: 2000-06-07
PRIOR APPLICATION NUMBER: USSN 60/221,836
PRIOR FILING DATE: 2000-07-28
PRIOR APPLICATION NUMBER: USSN 60/227,860
PRIOR FILING DATE: 2000-08-24
NUMBER OF SEQ ID NOS: 94
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 23
LENGTH: 831
TYPE: DNA
ORGANISM: Bacillus subtilis
FEATURE:
NAME/KEY: CDS
LOCATION: (1)..(831)
US-10-984-449-23

Query Match 49.4%; Score 244; DB 21; Length 831;
Best Local Similarity 100.0%; Pred. No. 4,16-64;
Matches 244; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 251 ATGAAAAAAGAACTGATTTTCTAAAAATGAAAGAGTCTGAAGAGCGATTGTCATGCTG 310
DB 1 ATGAAAAAAGAACTGATTTTCTAAAAATGAAAGAGTCTGAAGAGCGATTGTCATGCTG 60
QY 311 ACCGCTTATGATTTATCCGAGAGCTAACTGCTGAAGAGCGAGTCTGACATGATTTTA 370
DB 61 ACCGCTTATGATTTATCCGAGAGCTAACTGCTGAAGAGCGAGTCTGACATGATTTTA 120
QY 371 GTCGGTATTCATCTGGAATGCTGCTCGGCTTGAATCACTGCTGCTGACAGTT 430

Db 121 GTCGTGATTCATTGGAGATGCTCTCGGCTTGAATTCAGTGTGATGACAGTT 180
Qy 431 GCGGACATGATTCATCATACAAAAGCGTTTAAAGGGGTGGCGGAATACCTTTATGTG 490
Db 181 GCGGACATGATTCATCATACAAAAGCGTTTAAAGGGGTGGCGGAATACCTTTATGTG 240
Qy 491 ACAG 494
Db 241 ACAG 244

RESULT 9
US-10-508-768a-8/c
; Sequence 8, Application US/10508768A
; Publication No. US20050164335A1
; GENERAL INFORMATION:
; APPLICANT: WILLIAMS, MARK K.
; APPLICANT: YOCUM, R. ROGERS
; APPLICANT: PERO, JANICE G.
; TITLE OF INVENTION: METHODS AND ORGANISMS FOR PRODUCTION OF B6 VITAMERS
; FILE REFERENCE: OGI-002US
; CURRENT APPLICATION NUMBER: US/10/508,768A
; PRIOR FILING DATE: 2004-09-22
; PRIOR APPLICATION NUMBER: PCT/US03/008880
; PRIOR FILING DATE: 2003-03-21
; PRIOR APPLICATION NUMBER: 60/451,824
; PRIOR FILING DATE: 2003-03-03
; PRIOR APPLICATION NUMBER: 60/368,618
; PRIOR FILING DATE: 2002-03-29
; PRIOR APPLICATION NUMBER: 60/367,863
; PRIOR FILING DATE: 2002-03-25
; PRIOR APPLICATION NUMBER: 60/367,089
; PRIOR FILING DATE: 2002-03-22
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 3.3
; SEQ ID NO 8
; LENGTH: 10181
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: plasmid - pdx17R
US-10-508-768a-8

Query Match 45.5%; Score 224.6; DB 22; Length 10181;
Best Local Similarity 92.5%; Pred. No. 1.3e-57;
Matches 236; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

Qy 1 GCTAATGTTGGTCAAGCCCGTGTGATTTGTATCTTCATTGGGCAAGTATGCGCTG 60
Db 3570 GCTAATGTTGGTCAAGCCCGTGTGATTTGTATCTTCATTGGGCAAGTATGCGCTG 3511
Qy 61 CGAATCGACCTATTATTAATAATAGATAGACATTCAGAGATGCTGCTTGCATCAAAAA 120
Db 3510 CGAATCGACCTATTATTAATAATAGATAGACATTCAGAGATGCTGCTTGCATCAAAAA 3451
Qy 121 GCACTGGGACAGAGGATGAACTCGCCGAACTTTAGAAAGTAAAGATCTTCTGTTG 180
Db 3450 GCACTGGGACAGAGGATGAACTCGCCGAACTTTAGAAAGTAAAGATCTTCTGTTG 3391
Qy 181 TAAAGAAAGTTTTTGGCTTCAGAGAAAAAGGCAATCATCTCTCTTAACATGAGG 240
Db 3390 TAAAGAAAGTTTTTGGCTTCAGAGAAAAAGGCAATCATCTCTCTTAACATGAGG 3331
Qy 241 AGGAGAAAAATGAA 255
Db 3330 GTTATTAAGGCAATGAA 3316

RESULT 10
US-10-466-642-20/c
; Sequence 20, Application US/10466642
; Publication No. US20040048343A1
; GENERAL INFORMATION:

APPLICANT: OmniGene BioProducts
; TITLE OF INVENTION: METHODS AND MICROORGANISMS FOR THE PRODUCTION OF 3-(2-
; FILE REFERENCE: BGI-146PC
; CURRENT APPLICATION NUMBER: US/10/466,642
; CURRENT FILING DATE: 2003-07-18
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 20
; LENGTH: 6886
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: vector
; FEATURE:
; OTHER INFORMATION: PAN624
US-10-466-642-20

Query Match 43.8%; Score 216.2; DB 18; Length 6886;
Best Local Similarity 98.6%; Pred. No. 4e-55;
Matches 218; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 GCTAATGTTGGTCAAGCCCGTGTGATTTGTATCTTCATTGGGCAAGTATGCGCTG 60
Db 2909 GCTAATGTTGGTCAAGCCCGTGTGATTTGTATCTTCATTGGGCAAGTATGCGCTG 2850
Qy 61 CGAATCGACCTATTATTAATAATAGATAGACATTCAGAGATGCTGCTTGCATCAAAAA 120
Db 2849 CGAATCGACCTATTATTAATAATAGATAGACATTCAGAGATGCTGCTTGCATCAAAAA 2790
Qy 121 GCACTGGGACAGAGGATGAACTCGCCGAACTTTAGAAAGTAAAGATCTTCTGTTG 180
Db 2789 GCACTGGGACAGAGGATGAACTCGCCGAACTTTAGAAAGTAAAGATCTTCTGTTG 2730
Qy 181 TAAAGAAAGTTTTTGGCTTCAGAGAAAAAGGCAATC 221
Db 2729 TAAAGAAAGTTTTTGGCTTCAGAGAAAAAGGCAATC 2689

RESULT 11
US-10-466-717-20/c
; Sequence 20, Application US/10466717
; Publication No. US20040086982A1
; GENERAL INFORMATION:
; APPLICANT: Omogene Bioproducts
; TITLE OF INVENTION: PROCESSES FOR ENHANCED PRODUCTION OF PANTOTHENATE
; FILE REFERENCE: BGI-148PC
; CURRENT APPLICATION NUMBER: US/10/466,717
; CURRENT FILING DATE: 2003-07-18
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 20
; LENGTH: 6886
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: vector
; FEATURE:
; OTHER INFORMATION: PAN624
US-10-466-717-20

Query Match 43.8%; Score 216.2; DB 18; Length 6886;
Best Local Similarity 98.6%; Pred. No. 4e-55;
Matches 218; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 GCTAATGTTGGTCAAGCCCGTGTGATTTGTATCTTCATTGGGCAAGTATGCGCTG 60
Db 2909 GCTAATGTTGGTCAAGCCCGTGTGATTTGTATCTTCATTGGGCAAGTATGCGCTG 2850
Qy 61 CGAATCGACCTATTATTAATAATAGATAGACATTCAGAGATGCTGCTTGCATCAAAAA 120
Db 2849 CGAATCGACCTATTATTAATAATAGATAGACATTCAGAGATGCTGCTTGCATCAAAAA 2790

Qy 121 GGACTGGACAGAGGAGATGAATACTGCCGAACCTTAGAAGTAGAATCCTTCTCGTTG 180
| | | | |
Db 2789 GGACTGGACAGAGGAGATGAATACTGCCGAACCTTAGAAGTAGAATCCTTCTCGTTG 2730
| | | | |
Qy 181 TAAACGGAAGGTTTTTGGCTTGCAGAGAAAACGGCAATC 221
| | | | |
Db 2729 TAAACGGAAGGTTTTTGGCTTGCAGAGAAAACGGCAATC 2689
| | | | |

RESULT 12

US-10-956-157-121430
; Sequence 121430, Application US/10956157
; Publication No. US20050118625A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Mounts, William
; TITLE OF INVENTION: NUCLEIC ACID ARRAYS FOR DETECTING GENE EXPRESSION ASSOCIATED WITH
; FILE REFERENCE: 031896-043000 (AM 101081)
; CURRENT APPLICATION NUMBER: US/10/956,157
; CURRENT FILING DATE: 2004-10-04
; NUMBER OF SEQ ID NOS: 319805
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 121430
; TYPE: DNA
; ORGANISM: Control Sequence
US-10-956-157-121430

Query Match 41.1%; Score 203; DB 21; Length 6540;
Best Local Similarity 100.0%; Pred. No. 4,56-51;
Matches 203; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCTAATGTGTGGTACAAAGCCCGTGAATTTGGTATCTTCATTCATGGGCAATCGCCTG 60
| | | | |
Db 6338 GCTAATGTGTGGTACAAAGCCCGTGAATTTGGTATCTTCATTCATGGGCAATCGCCTG 6397
| | | | |
Qy 61 GGAATGCACTATTATTAAATATAGATGACATTCGACAGCTGCTGCTTGAATCCAAAAA 120
| | | | |
Db 6398 GGAATGCACTATTATTAAATATAGATGACATTCGACAGCTGCTGCTTGAATCCAAAAA 6457
| | | | |
Qy 121 GGACTGGACAGAGGAGATGAATACTGCCGAACCTTAGAAGTAGAATCCTTCTCGTTG 180
| | | | |
Db 6458 GGACTGGACAGAGGAGATGAATACTGCCGAACCTTAGAAGTAGAATCCTTCTCGTTG 6517
| | | | |
Qy 181 TAAACGGAAGGTTTTTGGCTTGC 203
| | | | |
Db 6518 TAAACGGAAGGTTTTTGGCTTGC 6540
| | | | |

RESULT 13

US-10-956-157-121431
; Sequence 121431, Application US/10956157
; Publication No. US20050118625A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Mounts, William
; TITLE OF INVENTION: NUCLEIC ACID ARRAYS FOR DETECTING GENE EXPRESSION ASSOCIATED WITH
; FILE REFERENCE: 031896-043000 (AM 101081)
; CURRENT APPLICATION NUMBER: US/10/956,157
; CURRENT FILING DATE: 2004-10-04
; NUMBER OF SEQ ID NOS: 319805
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 121431
; LENGTH: 6540
; TYPE: DNA
; ORGANISM: Control Sequence
US-10-956-157-121431

Query Match 41.1%; Score 203; DB 21; Length 6540;
Best Local Similarity 100.0%; Pred. No. 4,56-51;
Matches 203; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCTAATGTGTGGTACAAAGCCCGTGAATTTGGTATCTTCATTCATGGGCAATCGCCTG 60
| | | | |
Db 6338 GCTAATGTGTGGTACAAAGCCCGTGAATTTGGTATCTTCATTCATGGGCAATCGCCTG 6397
| | | | |
Qy 61 GGAATGCACTATTATTAAATATAGATGACATTCGACAGCTGCTGCTTGAATCCAAAAA 120
| | | | |
Db 6398 GGAATGCACTATTATTAAATATAGATGACATTCGACAGCTGCTGCTTGAATCCAAAAA 6457
| | | | |
Qy 121 GGACTGGACAGAGGAGATGAATACTGCCGAACCTTAGAAGTAGAATCCTTCTCGTTG 180
| | | | |
Db 6458 GGACTGGACAGAGGAGATGAATACTGCCGAACCTTAGAAGTAGAATCCTTCTCGTTG 6517
| | | | |
Qy 181 TAAACGGAAGGTTTTTGGCTTGC 203
| | | | |
Db 6518 TAAACGGAAGGTTTTTGGCTTGC 6540
| | | | |

RESULT 14

US-10-956-157-121432
; Sequence 121432, Application US/10956157
; Publication No. US20050118625A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Mounts, William
; TITLE OF INVENTION: NUCLEIC ACID ARRAYS FOR DETECTING GENE EXPRESSION ASSOCIATED WITH
; FILE REFERENCE: 031896-043000 (AM 101081)
; CURRENT APPLICATION NUMBER: US/10/956,157
; CURRENT FILING DATE: 2004-10-04
; NUMBER OF SEQ ID NOS: 319805
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 121432
; LENGTH: 6540
; TYPE: DNA
; ORGANISM: Control Sequence
US-10-956-157-121432

Query Match 41.1%; Score 203; DB 21; Length 6540;
Best Local Similarity 100.0%; Pred. No. 4,56-51;
Matches 203; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCTAATGTGTGGTACAAAGCCCGTGAATTTGGTATCTTCATTCATGGGCAATCGCCTG 60
| | | | |
Db 6338 GCTAATGTGTGGTACAAAGCCCGTGAATTTGGTATCTTCATTCATGGGCAATCGCCTG 6397
| | | | |
Qy 61 GGAATGCACTATTATTAAATATAGATGACATTCGACAGCTGCTGCTTGAATCCAAAAA 120
| | | | |
Db 6398 GGAATGCACTATTATTAAATATAGATGACATTCGACAGCTGCTGCTTGAATCCAAAAA 6457
| | | | |
Qy 121 GGACTGGACAGAGGAGATGAATACTGCCGAACCTTAGAAGTAGAATCCTTCTCGTTG 180
| | | | |
Db 6458 GGACTGGACAGAGGAGATGAATACTGCCGAACCTTAGAAGTAGAATCCTTCTCGTTG 6517
| | | | |
Qy 181 TAAACGGAAGGTTTTTGGCTTGC 203
| | | | |
Db 6518 TAAACGGAAGGTTTTTGGCTTGC 6540
| | | | |

RESULT 15

US-10-956-157-121456
; Sequence 121456, Application US/10956157
; Publication No. US20050118625A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Mounts, William
; TITLE OF INVENTION: NUCLEIC ACID ARRAYS FOR DETECTING GENE EXPRESSION ASSOCIATED WITH
; FILE REFERENCE: 031896-043000 (AM 101081)
; CURRENT APPLICATION NUMBER: US/10/956,157
; CURRENT FILING DATE: 2004-10-04
; NUMBER OF SEQ ID NOS: 319805
; SOFTWARE: PatentIn version 3.2

; SEQ ID NO 121456
; LENGTH: 6540
; TYPE: DNA
; ORGANISM: Control Sequence
US-10-956-157-121456

Query Match 41.1%; Score 203; DB 21; Length 6540;
Best Local Similarity 100.0%; Pred. No. 4.5e-51;
Matches 203; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
OY 1 GCTATGTTGGTACAGCCCGTTGATTTGGTATTAATTCATTGGGCAATATGCGCTG 60
Db 6338 GCTAATGTTGGTACAGCCCGTTGATTTGGTATTAATTCATTGGGCAATATGCGCTG 6397
OY 61 CGAATGCACTTATTATTAATAATAGATTCAGCAGTCTGCTTGATCCAAAAA 120
Db 6398 CGAATGCACTTATTATTAATAATAGATTCAGCAGTCTGCTTGATCCAAAAA 6457
OY 121 GCACTGGGACAGAGGATGAATCGCCGAATTTAGAAAGTGAAGATCCTTCGCTG 180
Db 6458 GCACTGGGACAGAGGATGAATCGCCGAATTTAGAAAGTGAAGATCCTTCGCTG 6517
OY 181 TAACGGAAGTTTTTGGCTTC 203
Db 6518 TAACGGAAGTTTTTGGCTTC 6540
```

Search completed: September 8, 2005, 23:56:05
Job time : 649 secs

This Page Blank (uspic)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 8, 2005, 16:54:28 ; Search time 3136 Seconds
(without alignments)
5996.093 Million cell updates/sec

Title: US-09-925-824A-1
Perfect score: 494
Sequence: 1 gcaaatgctgtgtacacg.....gaacaccttattgtgcag 494

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: EST:
1: gb_est1:
2: gb_est2:
3: gb_hic:
4: gb_est3:
5: gb_est4:
6: gb_est5:
7: gb_est6:
8: gb_gsa1:
9: gb_gsa2:

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	79.2	16.0	772	2	BE641179 Crl2.2.10
2	70.8	14.3	700	7	CV032340 RTNACL1.7
3	69.8	14.1	553	4	Bj178509 Bj178509
4	68.6	13.9	525	4	Bj413963 Bj413963
5	66.4	13.4	709	7	CK447171 N8A5.SP6
6	61	12.3	550	7	CF507610 USDA-FP-1
7	61	12.3	599	7	CF506499 USDA-FP-1
8	60	12.1	684	6	CB687940 CBEST-22-B
9	59.2	12.0	1159	4	BZ567907 pac82-164
10	58.8	11.9	879	4	Bj029615 Bj029615
11	58.6	11.9	505	8	BH236381 AU0CA58TF
12	58.6	11.9	1252	3	BH236381 AU0CA58TF
13	58.6	11.9	1678	3	CNS0A4X7 Arabidops
14	57.6	11.7	429	6	CA989682 EST643190
15	56.8	11.5	725	6	CD483545 acrol-26m
16	56.6	11.5	387	5	BQ493962 EST03128
17	56.6	11.5	447	5	BQ494730 EST03895
18	56.6	11.5	473	6	CA581877 EST001552
19	54.8	11.1	499	6	CB081675 hks5a09.5
20	54.8	11.1	585	6	BH398523 AG-ND-170
21	54.8	11.1	752	8	BH376769 AG-ND-134
22	53.8	11.0	703	5	BH065211 Pgr 6 P02
23	53.8	10.9	402	5	BU943310 Ujirmpst
24	53.6	10.9	1068	9	CNS06YTM AL421472 T3 end of

25	53	10.7	314	4	BG463405	BG463405 EM1_49 B0
26	53	10.7	481	4	BG411250	BG411250 EM1_27 E0
27	53	10.7	519	4	BG464232	BG464232 EM1_71 C0
28	53	10.7	528	2	Bf176896	Bf176896 EM1_4 A09
29	53	10.7	557	6	CD220973	CD220973 CCG1_72-B
30	53	10.7	656	6	CD224696	CD224696 CCG1_35-A
31	53	10.7	728	8	BH650859	BH650859 BOMCUB82TR
32	53	10.7	743	8	BZ448307	BZ448307 BONDN32TF
33	53	10.7	1200	9	CL958099	CL958099 Oa1FCC000
34	52	10.5	878	7	CO006055	CO006055 EST794390
35	51.8	10.5	578	1	AU235768	AU235768 AU235768
36	51.4	10.4	455	7	CF604653	CF604653 RADIC01_0
37	51.4	10.4	466	7	CF604658	CF604658 RADIC01_0
38	51.4	10.4	495	7	CF609965	CF609965 INF1001_0
39	51.4	10.4	661	8	AQ864621	AQ864621 n0b00023H
40	50.8	10.3	652	5	BQ471131	BQ471131 HV01D11T
41	50.8	10.3	653	1	AL504804	AL504804 AL504804
42	50.4	10.2	657	1	AL504805	AL504805 AL504805
43	50.2	10.2	416	8	BH410494	BH410494 1007018FO
44	50.2	10.2	447	4	BG275106	BG275106 WHE2117 H
45	50.2	10.2	538	7	CO525451	CO525451 3530_1_16

ALIGNMENTS

RESULT 1
LOCUS BE641179 772 bp mRNA linear EST 01-SEP-2000
DEFINITION Crl2.2.103 SP6 Ceratopteris Spore Library Ceratopteris richardii
CDNA-Clone-Crl2.2.103 5', mRNA sequence.
ACCESSION BE641179
VERSION BE641179.1 GI:9958840
KEYWORDS EST.
SOURCE Ceratopteris richardii
ORGANISM Ceratopteris richardii
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Monilliformes; Filicophyta; Filicopsida; Filicales; Pteridaceae;
Ceratopteris.

REFERENCE 1 (bases 1 to 772)
AUTHORS Chatterjee, A., San Miguel, P., Stout, S.C., Banks, J., and Roux, S.J.
TITLE Expressed sequence tags of cDNA clones from a C. richardii library
JOURNAL Unpublished (2000)
COMMENT Contact: Roux SJ
Section of Molecular Cell and Developmental Biology
University of Texas
Biology Building, Room 16, Austin, TX 78712, USA
Tel: 512 471 4238
Fax: 512 232 3402
Email: sroux@utcc.utexas.edu
Plate: Crl2.2 row: I column: 03
Seq primer: SP6.

FEATURES

source
1..772
/organism="Ceratopteris richardii"
/mol_type="mRNA"
/cuiovar="Brogna"
/db_xref="taxon:49495"
/clone="Crl2.2.103"
/tissue_type="Gametophyte"
/cell_type="Spore"
/dev_stage="20 hours after germination initiation"
/clone_lib="Ceratopteris Spore Library"
/note="Vector: pCMVSPORT6; EST sequence from cDNA library.
cDNA library constructed from mRNA isolated from C.
richardii spores that had developed for 20 hours after
their germination had been initiated by white light."

ORIGIN

Query Match 16.0%; Score 79.2; DB 2; Length 772;
Best Local Similarity 62.8%; Pred. No. 7.9e-12;
Matches 123; Conservative 0; Mismatches 73; Indels 0; Gaps 0;

FEATURES	source
Db	294 AACGATGTGATGCTGACCCGCTTATGATTTATCCGGCAGCTAACTTGCTGAACAAGCGG 355
Db	166 AACCATTATCAATGCTAAAGCGCTTATGACTATCTTCTGCTGTGTGATGTGCATAGCGG 225
Db	354 GAGTTCACATGATTTTATGCGGTATTCACCTTGGAATGGTGTCTCGCGCTTGATTCAA 413
Db	226 GCATTGACATCTGCTTTGTGCGAGATTTCTGTGGGTATGTATGTATGATCATATGATACCA 285
Db	414 CTGTGCGTGTGACAGCTTGCAGCATGATTCATACAAAAGCCGTTAAAGGGGTGGC 473
Db	286 CATTCACCTGTTACATGATGATATGCTGCTCCATTTGAGAGCAGTAGCAAGAGGTGCTA 345
Db	474 CGAATACCTTTATTTGT 489
Db	346 GACGATCTCTCTTGT 361
RESULT 2	
LOCUS	CV032340
DEFINITION	CV032340 700 bp mRNA linear EST 23-AUG-2004
ACCESSION	RTNACL1_7_A09.g2_A029 Roots plus added NaCl Pinus taeda cDNA clone
VERSION	CV032340
KEYWORDS	CV032340.1 GI:51495163
SOURCE	EST.
ORGANISM	Pinus taeda (loblolly pine)
TITLE	Pinus taeda
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Pinus; Pinus. 1 (bases 1 to 700)
AUTHORS	Pratt, L., Cordonnier-Pratt, M.-M., Lorenz, W. W., Zimmermann, C. and Dean, J. F. D.
COMMENT	An EST database from NaCl-treated loblolly pine (Pinus taeda) roots unpublished (2004) Other ESTs: RTNACL1_7_A09_b1_A029 Contact: Cordonnier-Pratt MM Laboratory for Genomics and Bioinformatics The University of Georgia, Department of Plant Biology Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA Tel: 706 542 1860 Fax: 706 583 0210 Email: mmpratt@uga.edu RNA prepared and library constructed by W. Walter Lorenz (School of Forest Resources, University of Georgia); plant material prepared by Craig Zimmermann (School of Forest Resources, University of Georgia) using rooted cuttings provided by the Forest Biology Research Cooperative (FBRC) and the CLONES project at the University of Florida; sequencing done in the Laboratory for Genomics and Bioinformatics, University of Georgia. Sequence ends have been trimmed to exclude vector and regions below phred quality 16. Three-prime sequences are presented as their reverse complement and have been trimmed to exclude polyA. Seq primer: JENREV (CAGGAACGCTATGACC). Location/Qualifiers 1..700 /organism="Pinus taeda" /mol_type="mRNA" /strain="3 CLONES" /db_xref="taxon:3152" /clone="RTNACL1_7_A09_A029" /lab_host="DH10B-T1 phage-resistant E. coli" /clone_11b="Roots plus added NaCl" /note="Organ: Root; Vector: pSL180; Site 1: EcoRI; Site 2: XhoI; The library was prepared from polyA+ RNA from the roots of 1-year-old loblolly pine (Pinus taeda) cuttings that were rooted and then planted in washed sand. The rooted cuttings were maintained for 135 days (July 2003 harvest) under ambient conditions in a local greenhouse. They were kept on a weekly regimen of 0.5x nutrient-complete Hoagland's solution and supplemented with additional water sufficient to maintain a 15% soil moisture content. Twenty-four hours (24h) prior to harvesting roots for mRNA preparation, the potted trees

[illegible]

ORIGIN

Qy	199	CTTGAGAAAGAAAGGAGATCATCTCTCTAAACATGAGAGAGAAACATGAAAC	258
Db	172	CCTGCTGAGACACGCTATATGAGAGCCCAAGTCAAGTCCCGTGGAAAGGGTAG	231
Qy	259	AAAACTGAGATTTCTAAAAATGAAAGAGTCTGAGAACCGATTTTCATGCTGACCGCTTA	318
Db	232	GTTGGGGACACTTGAAGCCAAGTATCAACAGATCAGCCCAACAGATGTTAACTGGCTA	291
Qy	319	TGATTATCCGGAGCTAAACTTGCTGAACAAGCGGAGTTGCATGATTTTATGTCGTTGA	378
Db	292	TGATTATCCCTCCGGGCGCATGTGATCGACAGCATGACATATGTCTGTAGGGGA	351
Qy	379	TTCACTTGAATGATGTCCTTCGGCCTTGATTCAACTGTGCGTGTGACAGTTTCGGACAT	438
Db	352	CTCAGTGGTATGTTGTGCAATGGGCATGACAAACGCTGCACATGATGAGGACAT	411
Qy	439	GATCATCATCAAAAGCCGTTAAAGGGGGTGC	471
Db	412	GCTGCTGATTCGAAGCGGTGACAAAGGGGCGC	444

FEATURES	Location/Qualifiers
source	1. .525

```
/organism="Dictyostelium discoideum"  
/mol_type="mRNA"  
/strain="AX4"
```

ORIGIN

Qy	Db
239	106
GGAGAGAAAAACATGAAAAACAAACTGATTTTCTAAAAAAAGAAAGCTGCGAAACCG	GGAGAAAATAAAAAAAAGACAGTTTGTGATATTTTAAAGAAATATTAAGATGATATACCA
299	166
ATTGTCACTGCACCGCTTATGATTATTCGGCAGCTAACTTCTGGAACAAGCGGAGATT	ATTTCATATGGTTACAGCATATGATATTTTAAAGATTCAAGATTAAGTTGATTAATCAGGAATG
359	226
GACATGATTTTAACTCGGAGATTCACTTGGATGCGTGGCTCTGATTTCAACTGTC	GATATGATCTCGTGGAGATTTCATTTAGGTATGTTATGAAATGATGAGAGGTACACA
419	286
GGTGTGACAGTTGCGGACATGATCATCATACAAAAGCGTTAAAAAGGGGTGCGCCGAAT	TCAGTGACAAATGAAACAATATGATTATCATTTGTAATACATGATGATGAAGGTAGTAAAAAG
479	346
ACCTTATTTGT 489	TCATTTTGTAGT 356

RESULT 5	709 bp	mRNA	linear	EST 15-JAN-2004
CK447171				
LOCUS	CK447171			
DEFINITION	NB85.SP6 Aspergillus nidulans negative subtraction hybridization			

ACCESSION	CK447171	GI:40884648
VERSION	CK447171.1	
KEYWORDS	BST.	
SOURCE	<i>Emericella nidulans</i> (anamorph: <i>Aspergillus nidulans</i>)	

Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes; Eurotiales; Trichocomaceae; Emmentella.

REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
1 (bases 1 to 709)	Ray, A., Macwana, S., Ayoubi, P., Hall, L. T., Prade, R. and Mort, A. J.	Negative subtraction hybridization: An efficient method to isolate large numbers of condition-specific cDNAs	BMC Genomics 5, 22 (2004)	Contact: Patricia Ayoubi

Department of Biochemistry and Molecular Biology
Oklahoma State University
248 Noble Research Center, Stillwater OK, 74078
Tel: 405-744-6209
Fax: 405-744-7799
Email: ayoub@okstate.edu
PCR Primers
FORWARD: SP6
BACKWARD: T7
Insert Length: 1 Std Error: 0.00
Seq primer: SP6.

FEATURES	Location/Qualifiers
source	1. .709

```

/organism="Emeticella nidulans"
/mol_type="rRNA"
/strain="FGSC C26"
/db_xref="taxon:162425"
/clone="N8A5"
/tissue_type="vegetative mycelia"
/lab_host="E. coli"

```


Db 150 AGAAGGAGAGCCATTCATATGTCAACCGCTATGATATCCCTCGGCGGTGACCTCG 209

Qy 345 AACAGCGGAGTTGACATGATTTTAACTGCTGATTCATTGAAATGCTGCTCGGCC 404

Db 210 AAGGTGGGGATGATATATATGTTGGTCGTGACTCAGCGGCATGCTGCTTCAAGCTC 269

Qy 405 TTGATTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 464

Db 270 ACGACATCTAGCTTCCCATCTCTCGAGAAATGCTTGTCTGCTGCTGCTGCTGCTGCT 329

Qy 465 GGGGTGGCGCGAATACCTTTATGT 489

Db 330 GTGGCGCTAAGACCGTTACTTGT 354

RESULT 8
CB687940
LOCUS
DEFINITION CB687940 684 bp mRNA linear EST 04-SEP-2003
CBST-22-B-10 Mixed source, strain EP155 and EP155 infected with hypovirus CHV1-EP713 Cryphonectria parasitica cDNA clone EP155, EP155-CHV1-EP713 5-prime, mRNA sequence.

ACCESSION
VERSION CB687940
KEYWORDS
SOURCE
ORGANISM
CRYPHONECTRIA PARASITICA
CRYPHONECTRIA PARASITICA
EUKARYOTA; FUNGI; ASCOMYCOTA; PEZIZOMYCOTINA; SORDARIOMYCETES; SORDARIOMYCETIDAE; DIAPORTHALES; VALSACEAE; CRYPHONECTRIA-ENDOTICHA COMPLEX; CRYPHONECTRIA.

REFERENCE
AUTHORS
TITLE
Dawe, A.L., McMains, V.C., Panglao, M., Kasahara, S., Chen, B. and Nuss, D.L.
An ordered collection of expressed sequences from Cryphonectria parasitica and evidence of genomic microsynteny with Neurospora crassa and Magnaporthe grisea
Microbiology 149 (9), 2373-2384 (2003)

JOURNAL
MEDLINE
PUBMED
COMMENT
Contact: Dave AL, Nuss DL
Center for Biosystems Research
University of Maryland Biotechnology Institute
5115 Plant Sciences Building, College Park, MD 20742, USA
Tel: 301-405-7661 or 5111
Fax: 301-314-9075
Email: dave@umbi.umd.edu
Cryphonectria parasitica EST derived from mixed samples of cDNA from wild-type (strain EP155) and hypovirus-infected cultures.
Seq primer: M13-reverse

FEATURES
source
1. 684
Location/Qualifiers
/organism="Cryphonectria parasitica"
/mol_type="mRNA"
/strain="EP155"
/db_xref="taxon:38755 and 52571"
/db_xref="taxon:5116"
/clone="EP155, EP155-CHV1-EP713"
/clone_1id="Mixed source, strain EP155 and EP155 infected with hypovirus CHV1-EP713"
/note="Cryphonectria parasitica EST derived from mixed samples of cDNA from wild-type (strain EP155) and hypovirus-infected cultures"

ORIGIN
Query Match 12.1%; Score 60; DB 6; Length 684;
Best Local Similarity 55.2%; Pred. No. 3.1e-06;
Matches 117; Conservative 0; Mismatches 95; Indels 0; Gaps 0;

Qy 278 ATGAAGAGTCTGAAGAACGATGCTGCTGACCGCTTATGATTTATCCGACCTTAA 337

Db 208 ATGTATCAAGAAACGAGCCATCAACGATGATCAAGCCCATGACTCCGCTCCCTGAC 267

Qy 338 CTTCGTGAACAAGCGGAGTTGATGATTTATGCTGGATTCCTTGATGATGCTGCTC 397

Db 268 GTTGGCGATGACGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 327

Qy 398 CTGGCTTATTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 457

Db 328 CTGGGATGAGAGACAAAGTGAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 387

Qy 458 GTTAAAGGGGTGCGCGAATACCTTTATGT 489

Db 388 GTTCCAGGCGACAGTCCGATTTACCGT 419

RESULT 9
B2567907/c
LOCUS
DEFINITION B2567907 1159 bp DNA linear GSS 17-DEC-2002
pac82-164_7275.y2 pac82-164 Pseudomonas aeruginosa genomic clone pac82-164_7275, genomic survey sequence.

ACCESSION
VERSION B2567907
KEYWORDS
SOURCE
ORGANISM
PSEUDOMONAS AERUGINOSA
PSEUDOMONAS AERUGINOSA
BACTERIA; PROTEOBACTERIA; GAMMAPROTEOBACTERIA; PSEUDOMONADES; PSEUDOMONADACEAE; PSEUDOMONAS.

REFERENCE
AUTHORS
TITLE
Spencer, D.H., Raymond, C.K., Smith, E.E., Sims, B.E., Hastings, M., Burns, J.L., Kaul, R. and Olsen, M.V.
Whole-Genome-Sequence variation among multiple isolates of Pseudomonas aeruginosa library
J. Bacteriol. (2002) in press
Contact: Chris K. Raymond
Genome Center
University of Washington
Box 352145, Seattle, WA 98105-2145, USA
Tel: 2062216954
Fax: 2066857244
Email: craymon@u.washington.edu
Class: shotgun.

FEATURES
source
1. 1159
Location/Qualifiers
/organism="Pseudomonas aeruginosa"
/mol_type="genomic DNA"
/strain="2-164"
/db_xref="taxon:287"
/clone="pac82-164_7275"
/clone_1id="pac82-164"
/note="clinical isolate 2-164 Whole genomic shotgun library."

ORIGIN
Query Match 12.0%; Score 59.2; DB 8; Length 1159;
Best Local Similarity 54.6%; Pred. No. 5.9e-06;
Matches 118; Conservative 0; Mismatches 98; Indels 0; Gaps 0;

Qy 279 TGAAGAGTCTGAAGAACGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 338

Db 350 TGAAGCAGAGGCGGAGAAAGTCCCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 291

Qy 339 TTGCTGAACAAGCGGAGTTGATGATTTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 398

Db 290 CCGGACGCAAGCGCGGCTGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 231

Qy 399 TCGGCTTGAATTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 458

Db 230 AGGCTCAAGACGACCTGCGGCTGACGAAAGATGCTTATCAACAGGCTTGC 171

Qy 459 TTAAGAGGGGTGCGCGAATACCTTTATTTGAGAG 494

Db 170 TGAAGCGGCTAACAAGGTTCCCTGATGCTGACG 135

RESULT 10
B029615/c
LOCUS
B029615 579 bp mRNA linear EST 26-SEP-2003

DEFINITION B029615 NIBB Mochii normalized Xenopus neurula library Xenopus laevis cDNA clone Xl012p21 5', mRNA sequence.

ACCESSION B029615

VERSION B029615.1

KEYWORDS GI:17369985

SOURCE EST.

ORGANISM Xenopus laevis (African clawed frog)

Xenopus laevis

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae; Xenopodinae; Xenopus; Xenopus.

1 (bases 1 to 579)

Kitayama,A., Terasaka,C., Mochii,M., Ueno,N., Shin-I,T. and Kohara,Y.

Expressed genes in X. laevis embryo

Unpublished (2001)

Contact: Tadasu Shin-i

Center For Genetic Resource Information

National Institute of Genetics

1111 Yata, Mishima, Shizuoka 411-8540, Japan

Tel: 81-559-81-6856

Fax: 81-559-81-6855

Email: tshin@genes.nig.ac.jp

The information of this clone is available through the following URL.

http://xenopus.nibb.ac.jp.

FEATURES

source

location/Qualifiers

1..579

/organism="Xenopus laevis"

/mol_type="mRNA"

/db_xref="taxon:8355"

/clone="Xl012p21"

/tissue_type="whole embryo"

/dev_stage="stage 15"

/clone_lib="NIBB Mochii normalized Xenopus neurula library"

ORIGIN

Query Match 11.9%; Score 58.8; DB 4; Length 579;

Best Local Similarity 54.7%; Pred. No. 6.6e-06;

Matches 117; Conservative 0; Mismatches 97; Indels 0; Gaps 0;

281 AAGAGTCTGAAGACCGATTGTGATGTCGACCGCTTATGATTCGCGACGCTAACTT 340

562 AAACGGAAGAAAAACGTTTCCGACCATCACCGTTTATGACTATAGCTTGCGCAATTC 503

341 GCTGACAAAGCGGAGTTGACATGATTTTATGTCGATTCACCTTGGATGCTCTCTC 400

502 TTTGCTGATGAAGGCGCTTAACGTCATGCTGCGGCGATTCCGCGCATGACGCTTCAG 443

401 GGCCTGATTCACGTCGCTGTCGACGTCGACGATGATCATCAAAAGCCGTT 460

442 GGGCAGCATCTCACCTCGCAGTTACCGTTGCCGATATGCTTACCACTGCGCCGCTA 383

461 AAAAGGGGTGCGCCGATACCTTTATTTGACAG 494

382 CGTCGCGGCGCAACCACTGCTGCTGCTG 349

RESULT 11

LOCUS BH236381/c 805 bp DNA linear GSS 13-NOV-2001

DEFINITION AUTGAS5TF AUTC Arabidopsis thaliana genomic clone AUTGAS5, genomic survey sequence.

ACCESSION BH236381

VERSION BH236381

KEYWORDS GI:16906739

SOURCE GSS.

ORGANISM Arabidopsis thaliana (thale cress)

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

1 (bases 1 to 805)

Town,C.D., Whitelaw,C.A., Pal,G., Van Aken,S.E., Utezbek,T.V.,

TITLE Feldjlym,T.V. and Fraser,C.M.

JOURNAL Survey sequencing of Arabidopsis thaliana BAC T1317

COMMENT Unpublished (2001)

Other_GSSs: AUTGAS5TR

Contact: Chris Town

TIGR

9712 Medical Center Drive, Rockville, MD 20850, USA.

Tel: 301-838-3523

Fax: 301-838-0208

Email: cdtown@tigr.org

From Wash. U contig 1154.

Seq primer: TF

Class: sheared ends.

FEATURES

source

location/Qualifiers

1..805

/organism="Arabidopsis thaliana"

/mol_type="genomic DNA"

/ecotype="Columbia"

/db_xref="taxon:3702"

/clone="AUTGAS5"

/clone_lib="AUTC"

/note="Vector: PHO52; Site 1: BstXI; 2-3 kb sheared BAC DNA inserted into PHO52 using BstXI linkers"

ORIGIN

Query Match 11.9%; Score 58.6; DB 8; Length 805;

Best Local Similarity 53.8%; Pred. No. 8.1e-06;

Matches 121; Conservative 0; Mismatches 104; Indels 0; Gaps 0;

265 GGAATTTCTAATAAGAGAGCTGTGAAGACCGATTCATGCTGACCGCTTATGATTA 324

380 GCAATTTAGACAGAAAGCATCGAAAGGTGAGCGATTAATGATGTCACCGCTACGATTA 321

325 TCCGCGCATTAACCTTCCTGAACAAGCGGAGTTGACATGATTTTGTGCGGTATTCAC 384

320 CCTTCGCGGTTTCACTCACTCACTGATGATGATGATGATGATGATGATGATGATGATG 261

385 TGAATGTCGTCCTCGGCTTCGATTCATGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 444

260 TGCATGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 201

445 TCATTAACAAAGCCGTTAAAGGGGTGCGCGAATACCTTTATTTGT 489

200 TCACTGTCGCGCGCTTCTCGCGAGCCAAAGACCACTTCTTGT 156

RESULT 12

LOCUS CNS0A78Y 1252 bp mRNA linear HTC 06-FEB-2004

DEFINITION CNS0A78Y Arabidopsis thaliana Full-length cDNA Complete sequence from clone GSRTP572FE06 of flowers and buds of strain col-0 of Arabidopsis thaliana (thale cress).

ACCESSION BX822672

VERSION BX822672.1

KEYWORDS GI:42466004

SOURCE HTC; GSI; cDNA.

ORGANISM Arabidopsis thaliana (thale cress)

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

1 (bases 1 to 1252)

Castelli,V., Aury,J.M., Jallion,O., Wincker,P., Clepet,C., Menard,M., Cruaud,C., Queier,F., Scarpelli,C., Schachter,V., Temple,G., Caboche,M., Weissenbach,J. and Salanoubat,M.

Whole Genome Sequence Comparisons and 'Full-length' cDNA Sequences: A Combined Approach to Evaluate and Improve Arabidopsis Genome Annotation

Unpublished

2 (bases 1 to 1252)

Genoscope.

Direct Submission

Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage :

BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr

COMMENT

- Web : www.genoscope.cns.fr
 The sequences are based on single pass reads.
 Life Technologies (a division of Invitrogen) members carried out full-length libraries construction : Temple G.
 Genoscope members carried out sequencing and annotation : Castellani V., Aury J.M., Jallion O., Wincker P., Menard M., Cruaud C., Schachter V., Weissenbach J., Salanoubat M.
 URGV INRA : Clepet C., Caboche M.
 Annotation is based on the June 2003 version of the Arabidopsis genome released by MIPS (Munich Information center for Protein Sequences) . 5 prime and 3 prime are assembled with Phrap.
http://www.genoscope.cns.fr/externe/sequences/Banque_Projet_EF/Full_length
<http://www.genoscope.cns.fr/cgi-bin/gdb/gdb?source=Arabidopsis.Location/Qualifiers>

FEATURES

source

1. .1252
 /organism="Arabidopsis thaliana"
 /mol_type="mRNA"
 /strain="Col-0"
 /db_xref="taxon:3702"
 /clone="GSLTFS572F06"
 /tissue_type="Flowers and buds"
 /plasmid="PCWSP08T 6"
 complement(1. .1252)
 /gene="At3g61530"

gene

ORIGIN

Query Match 11.9%; Score 58.6; DB 3; Length 1252;
 Best Local Similarity 53.8%; Pred. No. 9e-06;
 Matches 121; Conservative 0; Mismatches 104; Indels 0; Gaps 0;

265 GGATTTTCTAAATGAAGAGTGTGAAGAACCGATTGTCACTGACCGCTTATGATTA 324
 131 GCAATTGACAGAGATCGAAGAGTGAAGCCGATTAAGTGTACCGCTACGATTA 190
 325 TCCGGACGCTAACTGCTGTAACAGCGGAGTGTGACATGATTTAAGTCGGTATTC 384
 191 CCTTCCTCCCTTCAATGATGATGATGATGATGATGATGATGATGATGATGATG 250
 385 TGAATGCTGCTCGGCTTGAATCACTGCTGCTGATGACAGTTGGCGACATGATCA 444
 251 TGCATGTTGTCCATGCTATGACACCACTTCTCTATGAGGAGATGCTTGT 310
 445 TCATACAAAGCCGTTAAAGGGGTGCGCCGGAATACCTTTATTTGT 489
 311 TCATGTCGCGCGTGTCTCGCGAGCAAAAGACCACTTCTTGT 355

RESULT 13

CNS0A4X7 1678 bp mRNA linear HTC 06-FEB-2004
 LOCUS Arabidopsis thaliana Full-length cDNA Complete sequence from clone
 DEFINITION GSJLS20207 of Adult vegetative tissue of strain col-0 of
 Arabidopsis thaliana (thale cress).
 ACCESSION BX823298 GI:42462911
 VERSION BX823298.1
 KEYWORDS HTC; GSJLT cDNA.
 SOURCE Arabidopsis thaliana (thale cress)
 ORGANISM Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 1 (bases 1 to 1678)
 Castellani V., Aury J.M., Jallion O., Wincker P., Clepet C.,
 Menard M., Cruaud C., Queller P., Scarpelli C., Schachter V.,
 Temple G., Caboche M., Weissenbach J. and Salanoubat M.
 Whole Genome Sequence Comparisons and Full-length cDNA Sequences:
 A Combined Approach to Evaluate and Improve Arabidopsis Genome
 Annotation
 Unpublished
 2 (bases 1 to 1678)
 Genoscope.
 AUTHORS
 JOURNAL
 TITLE
 REFERENCE
 DIRECT Submission

JOURNAL

COMMENT

Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage :
 BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr)
 - Web : www.genoscope.cns.fr
 The sequences are based on single pass reads.
 Life Technologies (a division of Invitrogen) members carried out full-length libraries construction : Temple G.
 Genoscope members carried out sequencing and annotation : Castellani V., Aury J.M., Jallion O., Wincker P., Menard M., Cruaud C., Schachter V., Weissenbach J., Salanoubat M.
 URGV INRA : Clepet C., Caboche M.
 Annotation is based on the June 2003 version of the Arabidopsis genome released by MIPS (Munich Information center for Protein Sequences) . 5 prime and 3 prime are assembled with Phrap.
http://www.genoscope.cns.fr/externe/sequences/Banque_Projet_EF/Full_length
<http://www.genoscope.cns.fr/cgi-bin/gdb/gdb?source=Arabidopsis.Location/Qualifiers>

FEATURES

source

1. .1678
 /organism="Arabidopsis thaliana"
 /mol_type="mRNA"
 /strain="Col-0"
 /db_xref="taxon:3702"
 /clone="GSLTFS202D07"
 /tissue_type="Adult vegetative tissue"
 /plasmid="PCWSP08T 6"
 complement(1. .1678)
 /gene="At3g61530"

gene

ORIGIN

Query Match 11.9%; Score 58.6; DB 3; Length 1678;
 Best Local Similarity 53.8%; Pred. No. 9.7e-06;
 Matches 121; Conservative 0; Mismatches 104; Indels 0; Gaps 0;

265 GGATTTTCTAAATGAAGAGTGTGAAGAACCGATTGTCACTGACCGCTTATGATTA 324
 526 GCAATTGACAGAGATCGAAGAGTGAAGCCGATTAAGTGTACCGCTACGATTA 585
 325 TCCGGACGCTAACTGCTGTAACAGCGGAGTGTGACATGATTTAAGTCGGTATTC 384
 586 CCTTCCTCCCTTCAATGATGATGATGATGATGATGATGATGATGATGATGATG 645
 385 TGAATGCTGCTCGGCTTGAATCACTGCTGCTGATGACAGTTGGCGACATGATCA 444
 646 TGCATGTTGTCCATGCTATGACACCACTTCTCTATGAGGAGATGCTTGT 705
 445 TCATACAAAGCCGTTAAAGGGGTGCGCCGGAATACCTTTATTTGT 489
 706 TCATGTCGCGCGTGTCTCGCGAGCAAAAGACCACTTCTTGT 750

RESULT 14

CA989682 489 bp mRNA linear EST 06-JAN-2003
 LOCUS EST643190 Medicago truncatula cDNA clone GJLSD-42M4, mRNA
 DEFINITION sequence.
 ACCESSION CA989682 GI:27522576
 VERSION CA989682.1
 KEYWORDS EST.
 SOURCE Medicago truncatula (barrel medic)
 ORGANISM Medicago truncatula
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
 Medicago.
 1 (bases 1 to 489)
 Gruauk M.A., Sanac D., Town C.D., Van Aken S., Uterback T.,
 Cheung F. and Fraser C.M.
 ESTs from late stage developing seeds of Medicago truncatula
 Unpublished (2002)
 Contact: Gruauk M.A.
 USDA/ARS Children's Nutrition Research Center
 Baylor College of Medicine
 1100 Bates Street, Houston, TX 77030-2600, USA

Tel: 713 798 7044
Fax: 713 798 7078

Email: mgrusak@bcm.tmc.edu

TIGR sequence name: MTRCF7ATX

More information is available at: www.medicago.org

Seq primer: SKmod (CTA GAA CTA GCG GAT CC).

FEATURES

source
1. .489
/organism="Medicago truncatula"
/mol_type="mRNA"
/cultivar="A17"
/db_xref="taxon:3980"
/clone="pGLSD-42M4"
/tissue_type="Immature seeds"
/dev_stage="25 to 35 days after pollination"
/lab_host="XLOLR"
/clone_1ib="GLSD"
/note="Vector: pBluescript SK-; Site 1: EcoRI; Site 2: XhoI; Immature seeds, collected from pods ranging in age from 25 to 35 days after pollination, were harvested from greenhouse-grown plants. Seed were removed and separated from pod walls and immediately frozen in liquid nitrogen. Seeds throughout the age range were pooled for mRNA extraction. cDNA was prepared from polyA+ enriched RNA. The cDNA was directionally ligated into the Unizap XR vector from Stratagene and packaged using Gigapack III Gold packaging extracts. Plasmids containing cDNA inserts were excised from the recombinant lambda-zap phage using Ex-assist helper phage and propagated in XLOLR cells."

ORIGIN

Query Match 11.7%; Score 57.6; DB 6; Length 489;
Best Local Similarity 51.6%; Pred. No. 1.4e-05;
Matches 132; Conservative 0; Mismatches 124; Indels 0; Gaps 0;
QY 224 CATGAGAGGAGAAACATGAAACAACTGATTTCTTAAATGAAGAGTCTGAAG 293
DB 118 CAACATCAAAAAGAGTACACTTTCACACTGACAAACAAACATTAACCTCAC 177
QY 294 AACGATGTCATGTCGACCGCTTATGATTAATCCGACGCTAACTGCTGAACAGCG 353
DB 178 AGCCATACACATGCTCACTGCTTATGATTAACCTTCTGCTGTCACTCGATATGCGCG 237
QY 354 GAGTGACATGATTTTATGCGGTGATTCATCTGAAATGATGCTCGGCTGATTCAA 413
DB 238 CCATTGATATCTGTTGTTGATGATTCCTTCCATGGTGTATGATGATATCTA 297
QY 414 CTGTGCTGTGACAGTTCGACATGATTCATCAACAAACCGTTAAAGGGGTGCGC 473
DB 298 CTTTGCCATTAATCTTAAATGAAGCTGTTCAATTCGTCGTGCTGATGCTA 357
QY 474 CGAATACTTTATTTGT 489
DB 358 AAATCTCTCTCTGTGT 373

RESULT 15

CD483545

LOCUS CD483545 725 bp mRNA linear EST 04-JUN-2003

DEFINITION atr01-26ms1-d12 atr01 Amborella trichopoda cDNA clone

ACCESSION CD483545

VERSION CD483545.1 GI:31404813

KEYWORDS EST

SOURCE

ORGANISM

Amborella trichopoda

Amborella trichopoda

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; stem Magnoliophyta; Amborellaceae;

Amborella.

1 (bases 1 to 725)
dePamphilis,C., Soltis,D., Soltis,P., Farmerie,W., Ma,H.,
Tanksley,S., Leebens-Mack,J., Field,D., Buzgo,M., Kim,S., Ilut,D.,
Landherr,L., Hu,Y., Wall,K., Albert,V., Carlson,J., Doyle,J.,

TITLE

JOURNAL

COMMENT

Frohlich,M., Miller,W., Oppenheimer,D. and Theissen,G.
Generation of ESTs from early flower buds of Amborella trichopoda
Unpublished (2002)
Contact: Claude dePamphilis or James Leebens-Mack
Muller Laboratory
Penn State University
208 Mueller Laboratory, Department of Biology, ATTN Rm212, Penn
State University, University Park, PA 16802, USA
Tel: 814 863 6413
Fax: 814 865 9131

Email: cwd3@psu.edu or jh110@psu.edu

The sequence provided is trimmed of vector and low quality regions.

Full sequence and original trace file are available from the Plant

Genome Network website (<http://Pgn.cornell.edu>)

Plate: atr01-26ms1 row: d column: 12

Seq primer: M13P.

FEATURES

source
1. .725
/organism="Amborella trichopoda"
/mol_type="mRNA"
/db_xref="taxon:13313"
/clone="atr01-26ms1-d12"
/tissue_type="flower buds"
/dev_stage="2.5mm buds"
/lab_host="SOLR"
/clone_1ib="atr01"
/note="Vector: pBluescript SK (+/-); Site 1: EcoRI; Site 2: XhoI; Amborella trichopoda Ball; This library was made from male flowers only. Only floral buds with diameter of 2.5 mm or less were used for RNA isolation. This is a directionally cloned, non-normalized library. Avg. insert length: 1611; Primers: M13P and M13R; Antibiotic: 50 ug/ml Ampicillin; Primary Titer: 2.2486 pfu total; Amplified Titer: 1.37E10 pfu/ml; Mass Excised Titer: 3.53E9 total; This library has been generated by the Floral Genome Project (FGP). We would like to thank David Lorence at the National Tropical Botanical Garden for providing plant material for library building. The Floral Genome Project is funded by NSF's Plant Genome Research Program (DBI-0115684). More information about the project can be obtained at <http://fgp.bio.psu.edu>"

ORIGIN

Query Match 11.5%; Score 56.8; DB 6; Length 725;
Best Local Similarity 60.3%; Pred. No. 2.7e-05;
Matches 94; Conservative 0; Mismatches 62; Indels 0; Gaps 0;
QY 256 AACAAATCGATTTCTTAAATGAAGAGTGAAGAACGATGCTACGACCGC 315
DB 570 AACCTAAGAAATTTGGCGCTGAAGCATTAAGGGGAGACCTATACCATGTCACCTGC 629
QY 316 TTATGATTAATCCGACCTAACTTGCTGAACAAGCGGAGTGACATGATTTAGTCGG 375
DB 630 TTATGATTAATCCATGACCTGTTCACTCGATCAAGGAAATGATTAAGACTGTGG 689
QY 376 TGATTCATTGAATGTCGCTCGGCTTGATTC 411
DB 690 TGATTCGACGAGATGTCGTTCAATGGCATGATAC 725

Search completed: September 8, 2005, 22:05:28
Job time : 3143 secs